

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 06-30-03

Searcher: Besicly C4999

Terminal time: 20

Elapsed time: _____

CPU time: _____

Total time: 28

Number of Searches: _____

Number of Databases: 1

Search Site

____ STIC

____ CM-1

____ Pre-S

Type of Search

____ N.A. Sequence

____ A.A. Sequence

____ Structure

____ Bibliographic

Vendors

____ IG

____ STN

____ Dialog

____ APS

____ Geninfo

____ SDC

☒ DARC/Questel

____ Other CGN

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:50:54 ; Search time 1211.48 Seconds
(without alignments)
5454.265 Million cell updates/sec

Title: US-09-208-629F-6

Perfect score: 2157

Sequence: 1 CSMILQISXRLRDGTQVIM.....AYLTXXNDLREQQPSQRT 408

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101625_12930/app.query.fasta_1.1166
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09208629 @CGN 1 1 2013 @runat_24062003_101625_12930 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_mus:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	947.5	43.9	581	13	BM258063
2	790.5	36.6	554	13	BM255484
3	724	33.6	585	17	AZ820748
4	707	32.8	529	9	AA177828
5	615	28.5	495	9	AL713459
6	563.5	26.1	839	14	BQ897419
7	491.5	22.8	685	9	AL660446
8	476.5	22.1	938	17	CNS04548
9	464.5	21.5	765	13	BI183645
10	453.5	21.0	821	14	BM946931
11	437	20.3	801	13	BG924078
12	436.5	20.2	624	13	BI394593
13	423	19.6	860	9	AJ446632
14	422	19.6	800	9	AJ452624
15	421.5	19.5	808	9	AJ456719
16	418.5	19.4	890	9	AJ456561
17	416.5	19.3	741	9	AJ453402
18	413	19.1	265	10	BE146323
19	412.5	19.1	845	9	AJ456135
20	409.5	19.0	602	13	BM439733
21	409.5	19.0	747	9	AJ446611
22	409.5	19.0	773	9	AJ451396
23	408.5	18.9	749	9	AJ447002
24	405.5	18.8	766	9	AJ450835
25	401.5	18.6	740	9	AJ450921
26	401.5	18.6	760	9	AJ456491
27	400.5	18.6	730	9	AJ455912
28	391.5	18.2	817	17	CNS0303DLV
29	389.5	18.1	2020	11	BC013202
30	383.5	17.8	746	9	AJ450259
31	383.5	17.8	798	9	AJ451155
32	377.5	17.5	724	9	AJ447835
33	376	17.4	633	13	BJ068617
34	373.5	17.3	591	13	BI984999
35	372.5	17.3	682	9	AJ453748
36	370.5	17.2	668	10	BE005953
37	369.5	17.1	531	12	BF426487
38	368.5	17.1	1199	11	AF345566
39	364	16.9	904	9	AL532537
40	364	16.9	931	9	AL547762
41	364	16.9	1766	11	AK008952
42	363	16.8	2542	11	AK017378
43	363	16.8	3001	11	AK005013
44	362.5	16.8	746	9	AJ456692
45	361.5	16.8	542	9	AI942392

ALIGNMENTS

RESULT 1
LOCUS BM258063 581 bp mRNA linear EST 17-DEC-2001
DEFINITION 522739 MARC 3BOV Bos taurus cdna 5', mRNA sequence.
ACCESSION BM258063
VERSION BM258063.1 GI:17893662
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 581)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett

,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
Keel,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGAGC

Plate: 126 row: D column: 8

Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1. 581

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

BASE COUNT 122 a 171 c 117 g 171 t

ORIGIN

Alignment Scores:

Pred. No.: 6.11e-91 Length: 581

Score: 947.50 Matches: 176

Percent Similarity: 95.88% Conservaive: 10

Best Local Similarity: 90.72% Mismatches: 7

Query Match: 43.93% Indels: 1

DB: 13 Gaps: 1

US-09-208-629F-6 (1-408) x BM258063 (1-581)

Qy 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeuVal 123

Db 2 ATGGGTACCTGAGCAGCGCCCTTTAAGTACCAATTGATACCGCCATCTACATCTGGTG 61

Qy 124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 143

Db 62 TTTCAGTAGGTATGCGCGGCAATGCGGTGACCTGTGGATGCTC---TTCAGGACCCAGA 118

Qy 144 SerLeuCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163

Db 119 ACCATCCGTATGACCATCTCTACACCAACCTGGCCATTGCGACTTTCGTTTGTGT 178

Qy 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183

Db 179 ACATCGCCCTTTAGTAATAGCTTACCATCTCAATGGGAACAACCTGGGTATTTGGGAGGTC 238

Qy 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203

Db 239 ATGTGCGGGGACACCGGTGATCTTCTATGGCAACATGATGCTCATCTCTCTCTC 298

Qy 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223

Db 299 GCCTGCATCAGTATCAACCGGTACTTAGCCATTGTTTCATCTTTCACTTACCGGGAGCTG 358

Qy 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243

Db 359 CCCAAGCGCACCTACGCTTCTCATGTGGTGTGGCAACGCTTTCTTTATAC 418

Qy 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263

Db 419 ATGCTCGCGTTTTCATCTTGAAGCAGGAGTACTATCTTGTCCAGCAAGACATTACCACC 478

Qy 264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283

Db 479 TGCCACGATGTCACACACATCGAGTCTCTCGCCCTTCCAGCTCTACTACTTTCATC 538

Qy 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIle 297

Db 539 TCCTTGGCATTTTGGATTCTTAATCCCATTTCTGGTCATT 580

RESULT 2

BM255484 554 bp mRNA linear EST 17-DEC-2001

LOCUS 517232 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

DEFINITION BM255484

VERSION BM255484.1 GI:17891083

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 554)

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,

Casas,E., Wray,J.E., White,J., Cho,J., Fahrrenkrug,S.C., Bennett

,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,

Perteau,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and

Keel,J.W.

Sequence evaluation of four pooled-tissue normalized bovine cDNA

libraries and construction of a gene index for cattle

Genome Res. 11 (4), 626-630 (2001)

JOURNAL 21180013

MEDLINE

COMMENT

Contact: Smith TPL

USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 18

and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCAGTCACGAGC

Plate: 117 row: K column: 11

Seq primer: ATTAGGTGACACTATAG.

FEATURES

source

1. 554

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone_lib="MARC 3BOV"

/tissue_type="pooled"

/lab_host="DH10B"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from marrow, alveolar

macrophage, ovary, fetal semitendinosus muscle, and fetal

longissimus muscle."

BASE COUNT 140 a 148 c 118 g 148 t

ORIGIN

Alignment Scores:

Pred. No.: 3.41e-74 Length: 554

Score: 790.50 Matches: 151

Percent Similarity: 89.13% Conservaive: 13

Best Local Similarity: 82.07% Mismatches: 19

Query Match: 36.65% Indels: 1

DB: 13 Gaps: 1

US-09-208-629F-6 (1-408) x BM255484 (1-554)

Qy 17 ValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeuProThr 36

```

Db      6 CTGATCAAGATGAGACCGTTCATCTTTGACGCCATTGAGCACACTTCTCGCCCTGCT 65
Qy      37 PheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThrLeuProfil 56
Db      66 TCCTGTCAAGCGGCATGATATGATGATGACAGCAACTTGGCAAGCCAACTTGTCTATT 125
Qy      57 LysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGlu 76
Db      126 AAGACCTTCCTGGAGCTCCCAAAATTCCTTTGAAGAGTTCCTCCCTTCTGCCATAGAA 185
Qy      77 GlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHis 96
Db      186 GCCTGGACAGGACCAACAACTGTAATAATCAATGCCCTGAGAACCTTGATTCAAAT 245
Qy      97 LeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIle 116
Db      246 CTCCTATGTGAATAATGCTACCATGGGTACCTGAGCAGCCCTTTAAGTACCAATGATA 305
Qy      117 ProLalleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrp 136
Db      306 CCCGCCATCTACATCTCTGTTGTCAGTAGTATGCCGGCAATGCCGTGACCTGTGG 365
Qy      137 MetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIle 156
Db      366 ATGCTC---TTGAGGACAGAACCATCGGTATGACCATCTTCTACCAACCTGGCCATT 422
Qy      157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db      423 GCAGACTTCTGTTGTTGTACACTGCCCTTTAGATAGCTTACCATCTCAATGGGAAC 482
Qy      177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db      483 AACTGGGTATTTGGGAGGTCTATGTGCGGGCCACCAACGGTCACTCTTATGGCAACATG 542
Qy      197 TyrCysSerIle 200
Db      543 TATTGCTCCATT 554

```

RESULT 3

```

AZ820748
LOCUS      585 bp      DNA      linear      GSS 20-FEB-2001
DEFINITION 2M0093H12F Mouse 10kb plasmid UUGCIM library Mus musculus genomic
            clone UUGC2M0093H12 F, DNA sequence.

```

```

ACCESSION  AZ820748
VERSION     AZ820748.1
KEYWORDS    GSS
SOURCE      house mouse.

```

ORGANISM

```

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 585)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islam, H., Longacre, S., Mamoud, M., Meenen, E., Pedersen, T., Reilly
, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.
and Wright, D., Weiss, R.

```

REFERENCE

```

AUTHORS    Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            unpublished (2000)

```

TITLE

```

JOURNAL     Contact: Robert B. Weiss
COMMENT     University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunne@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0093 row: H column: 12
            Seq primer: CGTTGTAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 585.
            Location/Qualifiers

```

FEATURES

source

```

1. 585
/organism="Mus musculus"
/strain="C57BL/6J"
/db xref="taxon:10090"
/clone="UUGC2M0093H12"
/clone_lib="Mouse 10kb plasmid UUGCIM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pWB42 (G14732114|gb|AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 153 a 168 c 117 g 147 t
ORIGIN

```

Alignment Scores:

```

Pred. No.:      4,7e-67      Length:      585
Score:          724.00      Matches:    134
Percent Similarity: 82.90%      Conservative: 26
Best Local Similarity: 69.43%      Mismatches: 31
Query Match:    33.57%      Indels:     2
DB:             17          Gaps:         0

```

```

US-09-208-629F-6 (1-408) x AZ820748 (1-585)

```

```

Qy      44 AsnAspThrAsnLeuAlaLysProThrLeuProfilLysThrPheArgGlyAlaPro 63
Db      11 AATGTTTCAGACAACTCAGCAAGCCAACTTAACTATTAAAGATTTTAATGGGGTCCC 70
Qy      64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
Db      71 CAAATACCTTTGAGAAATTCACACTTCTGCATAGAGGGCTGGACAGGCCACACACA 130
Qy      84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
Db      131 ACTATAAAGCGGAGGTGTCGAGGACAGTATTTCACCTCTCCACGGTGAATAATGCTACC 190
Qy      104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuVal 123
Db      191 ATAGGATACCTGAGAAGTTCCTTAAGTACCAAGTATGATACCTGCATCTACATCTCTG 250
Qy      124 PheVal-ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrAr 143
Db      251 TTGTGAGTTGGTGTACCAAGCAACATCGTGACCTGTGGAACTCTCTTTAAGGACCAA 310
Qy      143 gSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVa 163
Db      311 ATCCATCAGTCTGTCTATCTTTTTCACCAACCTGGCCATCGCAGATCTCTTTTCTGT 370
Qy      163 lThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe-GlyGluV 183
Db      371 CACACTGCCATTTAAGATCGCTTCAATCGCAACCACTGGGTATTGGCGGAGG 430
Qy      183 alLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuL 203
Db      431 TCAGTGCCCGGATCACCACCGGTGTTTTCACCGCAACATGTACTGCGCTATCTGTATCC 490
Qy      203 euAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyL 223

```

```

Db      491 TCACCTGCGGGGATCAACCGTACTCGCCACGGCTCACCCCTTTCACATACAGNAGC 550
Qy      223 euProLysHisThrTyAlaLeuValThrCysGly 234
Db      551 TGCCCAAAAGCAGCTTCTCCATGCTCATGTGTGGC 585

RESULT 4
AA177828
LOCUS      529 bp mRNA linear EST 16-FEB-1997
DEFINITION 5' similar to TR:G663021 G663021 PROTEINASE ACTIVATED RECEPTOR 2.
            ; mRNA sequence.
ACCESSION AA177828
VERSION    AA177828.1 GI:1759090
KEYWORDS   EST.
SOURCE     house musculus.
ORGANISM   Mus musculus.
REFERENCE  1 (bases 1 to 529)
AUTHORS    Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
TITLE      The WashU-HMI Mouse EST Project
JOURNAL    Unpublished (1996)
COMMENT    Contact: Marra M/Mouse EST Project
            WashU-HMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:381203
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 494.
            Location/Qualifiers
                1..529
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="IMAGE:620379"
                /clone_lib="Soares mouse 3NBMS"
                /sex="male"
                /tissue_type="Spleen"
                /dev_stage="4 weeks"
                /lab_host="DH10B"
                /note="vector: pT7T3D-Pac (Pharmacia) with a modified
                polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
                was primed with a Not I - oligo(dT) primer [5',
                TGTTACCAATCTGAAGTGGAGCGCGCGCTGTTTTTTTTTTTTTTTTTTT
                3']; double-stranded cDNA was ligated to Eco RI adaptors
                (Pharmacia), digested with Not I and cloned into the Not I
                and Eco RI sites of the modified pT7T3 vector. RNA
                provided by Dr. Bertrand Jordan. Library went through
                three rounds of normalization, and was constructed by
                Bento Soares and M. Fatima Ronaldo."
BASE COUNT 110 a 168 c 107 g 144 t
ORIGIN

Alignment Scores:
Pred. No.:      2,67e-65      Length:      529
Score:          707.00      Matches:      125
Percent Similarity: 84.57%      Conservative: 23
Best Local Similarity: 71.43%      Mismatches: 27
Query Match:      32.78%      Indels:      0
DB:              9      Gaps:      0

US-09-208-629F-6 (1-408) x AA177828 (1-529)

```

```

Qy      109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyVal 128
Db      3 AGTTCTCTTAAGTACCACCAATGATACCTGCATCTACATCTGCTGTTGTGGTGTGTA 62
Qy      129 ProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThr 148
Db      63 CCAGCCCAACATCGTGACCTGTGGAAACTCTCTTAAGGACCAATCCATCAGTCTGGTC 122
Qy      149 ValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLys 168
Db      123 ATCTTTTACACCAACCTGGCCATCGCAGATCTCTCTTTCTGTGTACACATGCCATTAAG 182
Qy      169 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThr 188
Db      183 ATGCCTTACCATCTCAATGGCAACAACCTGGGTATTTGGCGAGGTACGTCGCCGATCACC 242
Qy      189 ThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIle 208
Db      243 ACGTTGTTTCTACGGCAACATGTAATGGCTATCTCTGATCTCCTACCTGCATGGGCATC 302
Qy      209 AsnArTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyr 228
Db      303 AACCGCTACTCGCCACGGCTCACCCCTTTCACATACAGAAAGCTGCCCAACGACGCTTC 362
Qy      229 AlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePhe 248
Db      363 TCCATGCTCATGTGTGGCATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 422
Qy      249 IleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHis 268
Db      423 ATCTGAGCAGAGGATACCACTCGTCCATCCGAGATCACCACCTGCCACGATGCTT 482
Qy      269 AsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
Db      483 GACCGCTGGCAGTCCCATCATCTCTCCGATTCTACTACTTCTGTC 527

RESULT 5
AA1713459
LOCUS      495 bp mRNA linear EST 22-MAR-2002
DEFINITION DKFZp686N1799 r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
            DKFZp686N1799 5', mRNA sequence.
ACCESSION  AL713459
VERSION     AL713459.1 GI:19696815
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
REFERENCE    1 (bases 1 to 495)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE        Wambutt,R., Heubner,D., Mewes,W., Weil,B. and Wiemann,S.
JOURNAL      EST (Wambutt,R., Heubner,D., Mewes,H.W., Weil,B. and Wiemann,S.)
COMMENT      Unpublished (1999)
            Contact: Wambutt R
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No sl sequence available.
            This clone (DKFZp686N1799) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
            Location/Qualifiers
                1..495
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="DKFZp686N1799"
                /clone_lib="686 (synonym: hlcc3)"
                /tissue_type="human skeletal muscle"
                /dev_stage="adult"
FEATURES
source

```

```

/lab_host="DH10B"
/notes=vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
BASE COUNT 132 a 128 c 107 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 1.65e-55 Length: 495
Score: 615.00 Matches: 121
Percent Similarity: 99.18% Conservative: 0
Best Local Similarity: 99.18% Mismatches: 1
Query Match: 28.51% Indels: 0
DB: 9 Gaps: 0

US-09-208-629F-6 (1-408) x AL713459 (1-495)
Qy 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
Db 129 TGTCCATGATTTTACAGATTTCATACGTTTAAAGAGCGGACTCAGTCAATAATG 188
Qy 21 LysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 189 AAGCCCTCATCTTTCGACGCTGCTGCTCTCTGTTGCTGCTGCTGCTGCTGCTGCTG 248
Qy 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
Db 249 GGCATGGAAATGATACAAACAACTTGGCAAGCCAACTTACCCATTAAGACCTTCGT 308
Qy 61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
Db 309 GGAGCTCCCAAAATCTTTTGAAGAGTTCCTCTTCTGCTTGGAGGCTGGACAGGA 368
Qy 81 AlaThrIleThrValIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
Db 369 GCCAGGATCTGTAATAAATTAAGTCCCTGAGAAAGTGCTTACATCTCCATGTGAA 428
Qy 101 AsnAlaThrMetGlyTyLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
Db 429 AATGCTACCATGGGTACCTGACCAGCTCTTAAAGTAACTAACTGATACCTGCCATCTAC 488
Qy 121 LeuLeu 122
Db 489 CTCCTG 494

RESULT 6
BO897419
LOCUS BO897419
DEFINITION AGNCOURT_8671134 NICHD XGC Emb4 xenopus laevis cDNA clone
IMAGE:4681483 5', mRNA sequence.
ACCESSION BO897419
VERSION BO897419.1 GI:22289433
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 839)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Plate: LLAM10561 row: k column: 20
High quality sequence stop: 737.

```

```

FEATURES
source
Location/Qualifiers
1..839
/organism="Xenopus laevis"
/db xref="taxon:8355"
/clone="IMAGE:4681483"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 2.1 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."
BASE COUNT 217 a 209 c 180 g 232 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.1e-49 Length: 839
Score: 563.50 Matches: 111
Percent Similarity: 62.08% Conservative: 38
Best Local Similarity: 46.25% Mismatches: 74
Query Match: 26.12% Indels: 17
DB: 14 Gaps: 4

US-09-208-629F-6 (1-408) x BQ897419 (1-839)
Qy 30 LeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAsp----- 45
Db 122 CTGTCTCTACTCTCTTTGTCGGGGCTGCTTAGGGCTCCAGATGACGATATGCTGGAG 181
Qy 46 ThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProPro--- 64
Db 182 AAAAATGAATCTGCTGTACCC-----AAGACATTCGGGGAAGAAAGATGCT 229
Qy 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThr----- 82
Db 230 GGGGACTATGAACAGCTTTCCCATAGCATTTATCCACGGCTTCAGAGACCACTACCAAC 289
Qy 83 -----IleThrValIleLysCysProGluGluSerAlaSerHis 96
Db 290 AACCATCTTCTTAAGAATCTCCAGCAATAATGCCTTTAAGATGCCAACAAAAACACC 349
Qy 97 LeuHisValLysAsnAlaThrMetGlyTyLeuThrSerSerLeuSerThrLysLeuIle 116
Db 350 ATCAAGTCAGCAACTTGACCTTGGGTACCTGAGTGGCAAGTAAGCAGGGCACTGATC 409
Qy 117 ProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrp 136
Db 410 CCAGGGCATTTATATATCATTTGTTGCTGCTGAGTACCAGCTAATGCCATTTGTTGG 469
Qy 137 MetLeuPhePheArgThrArgSerIleCysThrThrValPheThrAsnLeuAlaIle 156
Db 470 ATGTTGTTCCATCAGGTGACGCTGTGTGTCACAACTGTCTACGCCAGCTTAGCAACC 529
Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsn 176
Db 530 TCTGACTTGCTTATTGCTCATCTGCCTCATGCTCAAGATAGTTTATCATCTTAACCTGGAAC 589
Qy 177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 590 AACTGGATTTTGGAGAAACCATGTGCCGACTATGACATCTTCTTGTACTTTAAACATG 649
Qy 197 TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db 650 TACTGCTTATTTTGTCTCTGATGCTGTTAGCATCAGCGCTATGTTGCCATTTGTCAT 709
Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuLeuThrCysGlyLeuVal 236
Db 710 CCATTTATATACAGGAGCTTGCCTTANAGGACCTATACGGTGTCTTCTGCTGCTGCTG 769
Qy 237 TrpAlaThrValPheLeuTyrMetLeuProPheIleLeuLysGlnGluTyrTyrLeu 256
Db 770 TGTACCATTCGACTGGTCTACATGATCCCATTTTTTCATACACGGAACATATACATG 829

```

```

RESULT 7
AL660446      685 bp      mRNA      linear      EST 13-DEC-2001
LOCUS
DEFINITION    AL660446 XGC-neurula silurana tropicalis cDNA clone TNeu043b15 5',
              mRNA sequence.
ACCESSION    AL660446
VERSION      AL660446.1 GI:17674619
KEYWORDS     EST.
SOURCE       western clawed frog.
ORGANISM     Silurana tropicalis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
              Xenopodinae; Silurana.
REFERENCE    1 (bases 1 to 685)
AUTHORS      Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
TITLE        Sanger Xenopus tropicalis EST project 2001 (10/2001)
JOURNAL      Unpublished (2001)
COMMENT      Contact: Huckle E
              Sanger Centre
              Hinxton, Cambridgeshire, CB10 1SA, UK
              Email: trop@sanger.ac.uk
              Sanger Xenopus tropicalis EST project 2001
              TROPICALIS_SEQUENCE ID: TNeu043b15.sp6
              Sequencing primer: SP6
              This sequence is from a Xenopus Gene Collection (XGC) library
              constructed by Aaron M. Zorn.
FEATURES     source
              Location/Qualifiers
                1..685
                /organism="Silurana tropicalis"
                /db_xref="taxon:8364"
                /clone="TNeu043b15"
                /clone_lib="XGC-neurula"
                /dev_stage="neurula"
                /lab_host="Escherichia coli DH10B"
                /note="vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
                was oligo dt primed from sug of poly A+ RNA from neurula.
                EcoRI-NotI cut cDNA was then ligated into pCS107 with
                EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT   177 a 169 c 152 g 184 t 3 others
ORIGIN
Alignment Scores:
Pred. No.:    4.09e-42      Length:      685
Score:         491.50      Matches:    103
Percent Similarity: 61.09%      Conservative: 32
Best Local Similarity: 46.61%      Mismatches: 75
Query Match:   22.79%      Indels:     11
DB:            9           Gaps:       5

US-09-208-629F-6 (1-408) x AL660446 (1-685)
Qy 14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaGlyLeuLeuLeuVal 33
    |||
Db 26 GGNANAAGTACCACCTGAATGCAAGAGGCTGAAGATGAGTCCCTTCCTGCTATTCCT 85
    |||

Qy 34 LeuProThrPheCysGlnSerGlyMetClnAsnAspThrAsnLeuAlaLysProThr 53
    |||
Db 86 TTGTGTGGGTGCTCTGGGGCTCAAGATGAGGATGAGTCCCTGGAGAAAATGCTGCT 145
    |||

Qy 54 LeuProIleLysThrPheArgGlyAlaProPhe---AenSerPheGluGluPheProPhe 72
    |||
Db 146 GTCCCC---AGACATTTGGGGGAAGAAAGAGAGTGGGGAATATGAGAGCTTCCCAT 202
    |||

Qy 73 SerAlaLeuGluGlyTrpThrGlyAlaThrIleThr-----ValLysIleLys 88
    |||
Db 203 AGCATTTCTGCATGCGCTCCAGACACACACCAACCAATATCCTTTCTAAGAATCTCC 262
    |||

Qy 89 CysPro-----GluGluSerAlaSerHis-----LeuHisValLysAsnAlaThr 103
    |||
Db 263 AGCAATAATGCTTTAAAGTCCGACGCAACACACCATCATCAATAAGTCAGCAATTCACCC 322
    |||

Qy 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
    |||

```

```

Db 323 TTGGTGTACCTAAGTGGCAAAAGTAAGCAGGAACATGATACCGGCATCTATATCATTTGTT 382
    |||
Qy 124 PreValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheArgThrArg 143
    |||
Db 383 GTCTCATTTGGTGTGCGCAGCCATGCTTGGTACTATGATGCTGTTCATCAGGTGACA 442
    |||
Qy 144 SerIleCysThrThrValPheTyrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
    |||
Db 443 TCTGTGTGACCACTGTGCTTACGCCAGCTTACCCACTCTGACTTGTCTATTGCTCCTC 502
    |||
Qy 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
    |||
Db 503 ATGCTGCTCTTTAAGATAGCTTATCATCTTAAACGGCAACAACTGGATTTTGGAGAAACC 562
    |||
Qy 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeu 203
    |||
Db 563 ATGTGCGGGCCATGACCATTTTCTTGTACTTTTAAACATGTACTGCTCCATCTGCTCCTG 622
    |||
Qy 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
    |||
Db 623 ATGTGCTTTAGCATCAACCGCTAGTTCATTTGTCATTCATCCATTTATATACAGAGCTTG 682
    |||
Qy 224 Pro 224
    |||
Db 683 CCG 685

RESULT 8
CNS04SUP
LOCUS
DEFINITION    CNS04SUP 938 bp DNA linear GSS 24-MAY-2000
              Tetraodon nigroviridis genome survey sequence T7 end of clone
              001A08 of library H from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION    AL305674
VERSION      AL305674.1 GI:8199451
KEYWORDS     GSS; genome survey sequence.
SOURCE       Tetraodon nigroviridis.
ORGANISM     Tetraodon nigroviridis
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
              Tetraodontidae; Tetraodon.
REFERENCE    1 (bases 1 to 938)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,
              Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,
              Saurin, W. and Weissenbach, J.
TITLE        Human gene number estimate provided by genome wide analysis using
              Tetraodon nigroviridis DNA sequence
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 938)
AUTHORS      Roest-Crollius, H., Jaillon, O., Dasilva, C., Fizames, C., Fisher, C.,
              Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
              Weissenbach, J.
TITLE        Characterization and repeat analysis of the compact genome of the
              freshwater pufferfish Tetraodon nigroviridis
JOURNAL      Unpublished
REFERENCE    3 (bases 1 to 938)
AUTHORS      Genoscope.
TITLE        Direct Submission
JOURNAL      Submitted (12-APR-2000)
COMMENT      This sequence is a single read and was generated as part of a large
              scale clone-end sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/Tetraodon.
FEATURES     source
              Location/Qualifiers
                1..938
                /organism="Tetraodon nigroviridis"
                /db_xref="taxon:99883"
                /clone="001A08"
                /clone_lib="H"
                /notes="lib="H"
BASE COUNT   161 a 307 c 261 g 207 t 2 others
ORIGIN

```

Alignment Scores:

Pred. No.: 2,56e-40 Length: 938
 Score: 476.50 Matches: 97
 Percent Similarity: 58.30% Conservative: 47
 Best Local Similarity: 39.27% Mismatches: 92
 Query Match: 22.09% Indels: 12
 DB: 17 Gaps: 4

US-09-208-629F-6 (1-408) x CNS04SUP (1-938)

Qy 147 ThrThrValPheTyrThrAsnLeuAlaAlaPhePheCysValThrLeuPro 166
 Db 17 TCCTCCATCTTCATGGCAACCTGGCGTGGC-GACCTGCTCTTCGTCATCTGGGTCCG 75
 Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
 Db 76 CTGAATCGCTACCACTCAACGGGACACACTGGGTCTACGGGAGTCCCTGTGTGAAG 135
 Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle-206
 Db 136 GTCTGGTGGCGCTTCTTCTACGGCAACATGTAAGTCTCCATCGCTTCATCGCGTGCATC 195
 Qy 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
 Db 196 AGCTGTCCAGCGCTCAAGGCGCGTCTCCACCGCTCGACAGAGGGGG-----GGCAGC 249
 Qy 227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
 Db 250 GGTGGCGCGTGGCGGTGTCGCGCGCATCTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 309
 Qy 247 PhePheLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrTyrCysHisAsp 266
 Db 310 CTGTTCTGTACGATCAGCAGGTCTGTGTGACAAACCTGGGCATCCGCACTGCACGAC 369
 Qy 267 ValHisAsnThrCysGlySerSerProPheGlnLeuTyrTyrPheIleSerLeuAla 286
 Db 370 GTC-----ACCAGGCCCGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 423
 Qy 287 PhePheGlyLeuIleProPheValLeuIleTyrCysTyrAlaAlaIleArg 306
 Db 424 ACGTGGGCTTATGTTCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 483
 Qy 307 ThrLeuAsn-----AlaTyrAspHisArgTrpLeuTyrValLys 320
 Db 484 GCCCTCAGAAAGCATGACGGATCCGACCATCACCAGATGCGCGGAGCGCGTGGCC 543
 Qy 321 AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleLeu 340
 Db 544 TTGATTGTBTAGTCTGT 603
 Qy 341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
 Db 604 CTGTGGATTAATCTCTCTCTTTGAGCGGCTCTCCCAACACGCTGTACGGCGCTACGTC 663
 Qy 361 IleAlaLeuLysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPheLeuMet 380
 Db 664 ACCAGCTGTGCTGGCGGAGCTCAACAGCTGCTGGACCGCTTCTGTACTTCTATCTATC 723
 Qy 381 SerLysThr---ArgAsnHis 386
 Db 724 TCAGAGGACTTCAGGGAGCAC 744

RESULT 9

LOCUS B1183645 765 bp mRNA linear EST 10-JUL-2001
 DEFINITION UNL-P-FN-cb-d-02-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
 UNL-P-FN-cb-d-02-0-UNL 3', mRNA sequence.
 ACCESSION B1183645
 VERSION B1183645.1 GI:14658054
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 765)
 Caetano,A.R., Johnson,R.K. and Pomp,D.
 Generation and sequence characterization of a normalized cDNA
 library from swine ovarian follicles
 Unpublished (2001)
 Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message.
 Seq primer: M13 -29
 POLYA=No.

FEATURES
source

Location/Qualifiers
 1..765
 /organism="Sus scrofa"
 /strain="University of Nebraska, Lincoln Swine Selection
 Lines"
 /db_xref="taxon:9823"
 /clone="UNL-P-FN-cb-d-02-0-UNL"
 /clone_lib="UNL-P-FN"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: Eco RI; The UNL-P-FN
 library is a normalized library representing porcine
 ovarian follicles, ranging between 2.0 to 10.0 mm in
 diameter, collected during 7 days of the follicular phase
 of the pig estrous cycle. This library was derived from
 the library UNL-P-F2. The tag is a string of 5-6
 nucleotides present between the Not 1 site and the
 oligo-dT track. The library was constructed as described
 by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
 1996.

BASE COUNT 142 a 217 c 186 g 219 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 3.67e-39 Length: 765
 Score: 464.50 Matches: 99
 Percent Similarity: 54.90% Conservative: 41
 Best Local Similarity: 38.82% Mismatches: 94
 Query Match: 21.53% Indels: 21
 DB: 13 Gaps: 6

US-09-208-629F-6 (1-408) x B1183645 (1-765)

Qy 95 SerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys 114
 Db 9 TCAGGCTTTTCCTGGATGAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 68
 Qy 115 LeuIleProAlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaThr 134
 Db 69 TTTCTTCCTGGT 128
 Qy 135 LeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsn 153
 Db 129 CTGTGGGTCTTTCTTTTCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 188
 Qy 154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
 Db 189 CTGGCTTGGCGGACCT 248
 Qy 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
 Db 249 CATGGCAACAACCTGGGTTTACGGGGAGATCTCTTTGCAAGGTGTCTCATTTGCTTTTCTAC 308

Qy 194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
 Db 309 GGCAACATGTAAGTCTCATCTCTTCATGACCTGCTCAGCGTTCAGAGGTACTGGGTC 368
 Qy 214 IleValHisProPheThrTyrArgGlyLeuProIlyHisThrTyrAlaLeuValThrCys 233
 Db 369 ATTGTGAATCCATGGTGGTCAACCCAGGAGGAAAGCAAGTTCATCGCGCTCTCCCTG 428
 Qy 234 GlyLeuValTyrAlaThrValPheLeuTyrMetLeuProPhePheIleLeuValGlnGlu 253
 Db 429 GGA---ATATGGCTGCTGATTTCTGCTGTGACCATCCCTCCCTGTATGCTGGAACGACT 485
 Qy 254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
 Db 486 CTCTACATCCCGAGCCCTTCACATCACACCTGTCTAGCGTC----- 527
 Qy 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
 Db 528 ---TTGCCCGAGGAGGTGTTGGTAGGGGACATGTTCAATTATTTCCTCTCTCTGCCCATC 584
 Qy 288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
 Db 585 GGAGTCTTCCTGTTCCCGCCCTCTCTCAGGCCGCTGCTACGCTGCTGATGATCAGGAGC 644
 Qy 308 Leu-----AsnAlaTyrAspHis-----ArgTrpLeuTyrTyrValIleAla 321
 Db 645 CTGGGCTCTTCGTCATNGACGAAACTCAGGAAAGAGGAGCGGAGCCATTAGGCTC 704
 Qy 322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
 Db 705 ATCATCAGCGTCTGGCCATGATCTGATCTGCTTCAGCGCTAGC 749

RESULT 10

BM946931 821 bp mRNA linear EST 14-MAR-2002
 LOCUS UI-M-EH0P-bvn-k-05-0-UI.r1 NIH BMAP_EH0P Mus musculus cDNA clone
 DEFINITION IMAGE:5694460 5', mRNA sequence.

ACCESSION BM946931

VERSION BM946931.1 GI:19430516

KEYWORDS EST.

SOURCE house mouse.

ORGANISM

Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 821)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

CNA Library Preparation: Dr. M. Bento Soares, University of Iowa

CNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

FEATURES

source

Seq primer: pYX-5.

Location/Qualifiers

1..821

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5694460"

/clone_lib="NIH BMAP_EH0P"

/tissue_type="whole brain"

/dev_stage="embryo 18.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;

Site 2: Not I; The library was constructed according to

1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCACGAC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). Gene Discovery in the Developing Mouse Nervous System, supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 127 a 261 c 207 g 226 t

ORIGIN

Alignment Scores:

Pred. No.: 6,07e-38 Length: 821
 Score: 453.50 Matches: 99
 Percent Similarity: 57.39% Conservative: 64
 Best Local Similarity: 34.86% Mismatches: 90
 Query Match: 21.02% Indels: 32
 DB: 14 Gaps: 10

US-09-208-629F-6 (1-408) x BM946931 (1-821)

Qy 116 IleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsn-----AlaVal 133
 Db 3 ATGCCCTCCGTGACAGGATTTGTTTCATTGTTCAGCCTTCTCTGAACGCTCGGCCATC 62
 Qy 134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
 Db 63 GCAGTGTTCGTTTGGAGGATGAAGTCAAGAAG---CCGGCCGCTGGTGTACATGCTGCAC 119
 Qy 154 LeuAlaIleAlaAspPhePheCysValThrLeuProPheIleValIleAlaTyrHisLeu 173
 Db 120 CTGGCCATGCCGACGCTCTTCGTGTCGGTGTCTCCCTTCAAGATCAGCTACTACTTC 179
 Qy 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
 Db 180 TCCGGCACTGATGGCAGTTCGGGTCTGGATGTGGCGCTTCGCCACCGCAGCGTTTAC 239
 Qy 194 GlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
 Db 240 GGGAAACATGTACGCTCCATCATGCTCATGACGCTCATAGCATTTGACCGTTCCTGGCG 299
 Qy 214 IleValHisPro-----PheThrTyrArgGlyLeuProIlyHisThrTyrAlaLeu 230
 Db 300 GTGGTGTATCCGATCCAGTCCCTGCTCTGGCGCACTCTGGGCGGAGCAACTTC----- 353
 Qy 231 ValThrCysGlyLeuValTyrPalAlaThrValPheLeuTyrMetLeuProPhePheIleLeu 250
 Db 354 ---ACTTGGCTGCTCATTTGGGTGATGGCCATCATGGGGGTGGTGGCCCTCTCTCTCAAG 410
 Qy 251 LysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal-----His 268
 Db 411 GAGCAGACACCCCGAGTTCGGGACTCAACATCACCACCTGCCACGACGCTCCTCAGTCAG 470
 Qy 269 AsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePhe 288
 Db 471 AACCTGATGCAAGGC-----TTTACTGCTACTTCTCTGGGCTTCTTCGGCCATC 521
 Qy 289 GlyPheLeuIleProPheValLeuIleTyrCysTyrAlaAlaIleIleArgThrLeu 308
 Db 522 TTCTTTCTTGTGGCTGTGATCGTTTCCACCGTCTGTACACGCTCCATCATCCGGTGGCTG 581
 Qy 309 Asn-----AlaTyrAspHisArgTrpLeuTyrTyrValIleAla 321
 Db 582 AGCTCTCCCGGGTTCGCAACCGGAGCAAGAAGTTCGGGGCTTTGTTC-----CTG 632
 Qy 322 SerLeuLeuIleValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 341
 Db 633 TCTGCGCGGTGTGTGTCATCTTTCATCGTCTGTCTTGGGCCCCC---AACGCTCTCTCTATT 691

JOURNAL
COMMENT

Unpublished (2001)
Contact: Larry A. Coghurn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: coghurn@udel.edu, www.chickest.udel.edu.

FEATURES

source

1..624
/organism="Gallus gallus"
/strain="Commercial broiler chicken"
/db_xref="taxon:9031"
/clone="pgpln.pk002.112"
/clone.lib="Normalized Chicken
Pituitary/Hypothalamus/Pineal Library"
/sex="Male and female"
/tissue types="Pituitary Gland/Hypothalamus/Pineal Gland"
/dev stage="Embryonic (d12,d14,d19); post-hatch (w1,w3,w5,w7,w9)"
/lab host="E. Coli EMDH10B"
/notes="Vector: pCMVSPORT6; Library made from equivalent
pools of total RNA isolated from each tissue at different
ages. Single pass sequencing from 5'-end"

BASE COUNT 98 a 202 c 151 g 162 t 11 others
ORIGIN

Alignment Scores:

Pred. No.: 2.7e-36 Length: 624
Score: 436.50 Matches: 88
Percent Similarity: 58.79% Conservative: 29
Best Local Similarity: 44.22% Mismatches: 81
Query Match: 20.24% Indels: 2
DB: 13 Gaps: 1

US-09-208-629F-6 (1-408) x BI394593 (1-624)

Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db 25 CTCGTCTTCCTGGGCTGCCCTCCACGGCCGTG-GCCCTCTGGGTTCGGTACCAGG 83
Qy 142 ThrArgSerIleCysThrThrValPheThrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db 84 ACTGAGAAGCTGACTCCACCACCTCTCTGTATGAACCTGGCTGCAGCAGACCTGCTGCT 143
Qy 162 CysValThrLeuProPheLysIleAlaThrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db 144 ATCTCAGTGGTCCCTCAAGATATCTACTATTCTTGGGAACAACCTGCCCTTTGGG 203
Qy 182 GluValLeuCysArgAlaThrThrValIlePheThrGlyAsnMetTyrCysSerIleLeu 201
Db 204 GAAGTCTGTCCGGGTCCACACAGCTTCTTCTATGGGAACATGTACTGCTCAGTCTG 263
Qy 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db 264 CTGCTCACCTCATCAGTGTGGACGATACCTGGCGGTGGTGCATCTCTTCTCTCGCGC 323
Qy 222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db 324 TCTTTCCGTACCCCACTTTGCTGCCTGGACCTGTGCTGTCTGTGCTGTGCTGCC 383
Qy 242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 394 GGCATCACTGTCCCTGACTCTGCAGCAGCGTGTACCCCTGTACAGCAGACATC 443
Qy 262 ThrThrCysHisAspValHisAsnThrCysGlnSerSerSerProPheGlnLeuTyrTyr 281
Db 444 ACTCTGTCCACGATGTTCTCTCCAGCAGGAGCAGCAGCGG---TATTACTTCTACTAC 500
Qy 282 PheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyr 301
Db 501 TTCGTCTGCCTCATCGCTGGCTGCTTCTCTCTCTGTCAGTGTGCTGCTGCTGCTGCT 560
Qy 302 AlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTyrLeuTyrTyrValLys 320

Db 561 TGCTCAGTGTCTACAAGCCCNNNNNNNNCCAGNGGAGCGGTACTTCTNACTCCATGAAG 617

RESULT 13
AJ446632

LOCUS
DEFINITION
AJ446632
ACCESSION
VERSION
AJ446632.1 GI:20213853
KEYWORDS
EST.
SOURCE
ORGANISM

Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 860)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Buerstedde, J.M.

Gallus gallus bursal lymphocyte EST

Unpublished (2002)

Contact: Buerstedde JM

Cellular Immunology

Heinrich-Pette-Institute

Martinistr. 52, 20251 Hamburg, Germany

Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

FEATURES

source

1..860
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="14d16r1"
/clone.lib="rikeni"
/cell_type="bursal lymphocyte"
/dev stage="2-3 weeks old"
/note="CB inbred strain"

BASE COUNT 207 a 185 c 204 g 260 t 4 others
ORIGIN

Alignment Scores:

Pred. No.: 1.18e-34 Length: 860
Score: 423.00 Matches: 94
Percent Similarity: 54.69% Conservative: 40
Best Local Similarity: 38.37% Mismatches: 97
Query Match: 19.61% Indels: 15
DB: 9 Gaps: 6

US-09-208-629F-6 (1-408) x AJ446632 (1-860)

Qy 62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 96 GCCGCTCCCAACAGACAGTAGAACCCAGCAGTTCACAAAGAGAGAGAGAGAGAGAG 155
Qy 82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 156 AAGGTTCCAGATCTAGTAACGGCTCTGAGGAGTCA-----TACAAGTGGATGAC 206
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db 207 TTTGCAGCAAAAGCTCTTACAGGAAACACTACAGTTTTTCTCCACATGCTATGTC 266
Qy 122 LeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db 267 ATTGCTTTTATCTGTTGTTTCCCAAGCAATGCTATGCCCATCTGGGCTTTTTTTTTCAGA 326
Qy 142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db 327 ACAAGAAGAAACATCCTGCTGTGATTATATGTTAACTTGGCATTGGCAGACCTTCTC 386
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db 387 TTGTTGCTGCTGCTGCCACTCAAGATGTCATATTAATGGCAATAACTGGCTGTTT 446
Qy 181 GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 447 GGGGAAGGTCTCTGCAAAAGTCTTGTGGATTTTTTTATGAAATATGTACTGTCTCCATT 506

BASE COUNT 94 a 41 c 70 g 60 t
ORIGIN

Alignment Scores:
Pred. No.: 2.61e-34 Length: 265
Score: 413.00 Matches: 80
Percent Similarity: 96.39% Conservative: 0
Best Local Similarity: 96.39% Mismatches: 2
Query Match: 19.15% Indels: 1
DB: 10 Gaps: 0

US-09-208-629F-6 (1-408) x BE146323 (1-265)

Qy 273 SerSerSerProPheGlnLeuTyrPheIleSerLeuAlaPhePheGlyPheLeu 292
Db 248 TCCTCATCTCCCTCCCAACTCTATTACTTCATCTCCCTGGCGTTCTTTGAATTTAAAT 189
Qy 293 ProPheValLeuIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp 312
Db 188 CCATTGTGCTATCACTACTGCTATGCAGCCATCATCCGACACTTAATGATACAT 129
Qy 313 HisArgTrpLeuTyrValIleAlaSerLeuLeuIleValIlePheThrIleCys 332
Db 128 CATAGATGGTTGTGTATGTTAAGCGAGTCTCTCATCTTGTGATTTTACCATTGC 69
Qy 333 PheAlaProSerAsnIleIleLeuIleIleHisAlaSerTyr-TyrTyrAsnAnTh 352
Db 68 TGTGCTCCAGCAATATTATTCTTATTATTCACCATGCTACTATCTACTACACAC 9
Qy 352 rAspGly 354
Db 8 TCATGGC 2

RESULT 19
AJ456135 845 bp mRNA linear EST 22-APR-2002
LOCUS
DEFINITION AJ456135 riken1 Gallus gallus cdna clone 7k4r1, mRNA sequence.
ACCESSION AJ456135
VERSION AJ456135.1 GI:20266231
KEYWORDS EST.
SOURCE chicken.

ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Buerstedde, J.M.
Contact: Buerstedde JM
Unpublished (2002)
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.
Location/Qualifiers

FEATURES
source
1..845
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="7k4r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stages="2-3 weeks old"
/note="CB inbred strain"
194 a 185 c 213 g 252 t 1 others

BASE COUNT 194 a 185 c 213 g 252 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 1.52e-33 Length: 845
Score: 412.50 Matches: 84
Percent Similarity: 56.25% Conservative: 33
Best Local Similarity: 40.38% Mismatches: 85
Query Match: 19.12% Indels: 6

DB: 9 Gaps: 3
US-09-208-629F-6 (1-408) x AJ456135 (1-845)

Qy 62 AlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 133 GCCGCTCCACAGAGACAGTAGAACAGTTCAAAGGAGAAAGATTGTGTGGCCAG 192
Qy 82 ThrIleThrValIleIleCysProGluGluSerAlaSerHisLeuHisValIleAsn 101
Db 193 AAGGTTCCAGATACTAGTAACGCCTCTGAGGAGTCA-----TACAAGTGGATGAC 243
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db 244 TTTCGACGAAAGTCTTACAGGAAACTAATACAGATTTTCTTCCCATGCTATGTC 303
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db 304 ATTGTCTTTATCATTTGGTTGCCAAGCAATGCTATGCCCATCTGGGTCTTTTTTTTTCAGA 363
Qy 142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db 364 ACAAGAAGAAACATCCTGCTGTGATTTATATGTTAATGCTTGGCATTTGGCAGACCTTCTC 423
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db 424 TTGCTTGTCTGGTTCCTCACTGAAGATGATATCATTTAAATGCGAATAACTGGCTGTTT 483
Qy 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 484 GGGGAAGGCTCTGCAAGTGCTTGTGGATTTTTTTTATGGAATATGTTACTGCTCCATT 543
Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPropheThrTyr 220
Db 544 CTTTATTGACATGTCTCAGTGTGCAAGGATTTGGTGTGATGAGAACCCCATAGTGCAC 603
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrVal 240
Db 604 ---TCAAGAAGGAGTCTGAAATGCGCTGGGCATCTCCCTTCTATCTGGTACTGATT 660
Qy 241 Phe-LeuTyrMetLeuPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAs 260
Db 661 TTTGTTGGGAACCATTCGCTGTATCTTGTCAATCAGACNGTGTATATTTCAGACCTTAA 720
Qy 260 PileThrThrCysHisAspVal 267
Db 721 CATCACTACCTGCCATGATGTG 742

RESULT 20
BM439733/c
LOCUS
DEFINITION BM439733 602 bp mRNA linear EST 01-FEB-2002
pgrin.pk001.b9 Normalized Chicken Reproductive Tract cDNA Library
(pgrin) Gallus gallus cdna clone pgrin.pk001.b9 5' similar to
gb|AAK61908.1|AF384819.1 (AF384819) coagulation factor II
receptor-like 3 [Homo sapiens], mRNA sequence.
ACCESSION BM439733
VERSION BM439733.1 GI:18470508
KEYWORDS EST.
SOURCE chicken.
ORGANISM
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Cogburn, L.A. and Nys, Y.
ESTs from Normalized Chicken Reproductive Tract cDNA library-
University of Delaware and INRA, Tours-Poultry Unit Project
Unpublished (2002)
Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822


```
DB: 9 Gaps: 3
US-09-208-629F-6 (1-408) x AJ447002 (1-749)
Qy 62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 118 GCCGCTCCACAGACAGTAGAACGACGAGTTCAAAGGAAGAAAGTTTGTGGCCAG 177
Qy 82 ThrIleThrValIleLysCysProGluSerAlaSerHisLeuHisValLysAsn 101
Db 178 AAGGTTCAGACTAGTAACGCTCTGAGGAGTCA-----TACAAGTGGATGAC 228
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerHisLeuHisValLysAsn 121
Db 229 TTTGCAGCAAAAGTCTTACAGGAAACTAATCTAGCTTTTCTCCCACTGCTATGTC 288
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetPheArg 141
Db 289 ATTGCTTTTATCATGTTTCCCAAGCAATGCTATGCTGCTGGTCTTTTTCGAGA 348
Qy 142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db 349 ACAAGAAGAAACATCTGATGATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 228
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpValPhe 180
Db 409 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 288
Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPheThrTyr 220
Db 469 GGGGAAGTCTCTGCAAGTCTGCAAGGATTTGGGTGAGTGAACCCATAGTGCAC 588
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 589 ---TCAAGAAGAAAGTCTGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
Qy 241 PheLeuTyrMetLeuProPhePheLeuLeuGlnGluTyrTyrLeuValGlnProAsp 260
Db 646 TTGTTGGGAACCATTCGCTGATCTTGTCAATCAGACGNGTGATATTTTCAGACCTTAA 705
Qy 260 pileThrThrCysHis 265
Db 706 CATCACTACCTGCCAT 721
RESULT 24
AJ450835 766 bp mRNA linear EST 19-APR-2002
LOCUS riken1 Gallus gallus cdna clone 26115r1, mRNA sequence.
DEFINITION
ACCESSION AJ450835
VERSION AJ450835.1 GI:20218056
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
Buerstedde, J.M.
Contact: Buerstedde JM
Unpublished (2002)
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
Location/Qualifiers
1..766
/organism="Gallus gallus"
/db_xref="taxon:9031"
```

```
/clone="26115r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/notes="CB inbred strain"
BASE COUNT 181 a 160 c 197 g 227 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7,41e-33 Length: 766
Score: 405.50 Matches: 84
Percent Similarity: 56.52% Conservative: 33
Best Local Similarity: 40.58% Mismatches: 85
Query Match: 18.80% Indels: 6
DB: 9 Gaps: 3
US-09-208-629F-6 (1-408) x AJ450835 (1-766)
Qy 62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 118 GCCGCTCCACAGACAGTAGAACGACGAGTTCAAAGGAAGAAAGTTTGTGGCCAG 177
Qy 82 ThrIleThrValIleLysCysProGluSerAlaSerHisLeuHisValLysAsn 101
Db 178 AAGGTTCAGACTAGTAACGCTCTGAGGAGTCA-----TACAAGTGGATGAC 228
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerHisLeuHisValLysAsn 121
Db 229 TTTGCAGCAAAAGTCTTACAGGAAACTAATCTAGCTTTTCTCCCACTGCTATGTC 288
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetPheArg 141
Db 289 ATTGCTTTTATCATGTTTCCCAAGCAATGCTATGCTGCTGGTCTTTTTCGAGA 348
Qy 142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db 349 ACAAGAAGAAACATCTGCTGCTGATTTATGCTGCTGCTGCTGCTGCTGCTGCT 408
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpValPhe 180
Db 409 TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
Qy 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 469 GGGGAAGTCTCTGCAAGTCTTGTGGATTTTATGCAATATGCTGCTGCTGCTGCT 528
Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPheThrTyr 220
Db 529 CTTTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 589 ---TCAAGAAGAAAGTCTGAAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
Qy 241 PheLeuTyrMetLeuProPhePheLeuLeuGlnGluTyrTyrLeuValGlnProAsp 260
Db 646 TTGTTGGGAACCATTCGCTGATCTTGTCAATCAGACGNGTGATATTTTCAGACCTTAA 704
Qy 261 IleThrThrCysHisaspVal 267
Db 705 ATCACTACCTGCCATGATGTG 725
RESULT 25
AJ450921 740 bp mRNA linear EST 19-APR-2002
LOCUS riken1 Gallus gallus cdna clone 27asr1, mRNA sequence.
DEFINITION
ACCESSION AJ450921
VERSION AJ450921.1 GI:20218142
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
```


Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

FEATURES
 . source
 Location/Qualifiers
 1..798
 /organism="Gallus gallus"
 /db_xref="taxon:9031"
 /clone="27m8r1"
 /clone.lib="riken1"
 /cell_type="bursal lymphocyte"
 /dev_stages="2-3 weeks old"
 /note="CB inbred strain"
 /note="CB inbred strain"
 BASE COUNT 187 a 170 c 199 g 236 t 6 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.76e-30 Length: 798
 Score: 383.50 Matches: 81
 Percent Similarity: 55.02% Conservative: 37
 Best Local Similarity: 38.76% Mismatches: 84
 Query Match: 17.78% Indels: 7
 DB: 9 Gaps: 3

US-09-208-629F-6 (1-408) x AJ451155 (1-798)

Qy	62	AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTTPThrGlyAla	81
Db	115	GGCGTCCACACAGACACAGTAGACACGACGTTCAAAAGGAGAAGTTTTGTGGCCGAC	174
Qy	82	ThrIleThrValIysIleLysCysProGluGluSerAlaSerHisLeuHisValIysAsn	101
Db	175	AAGGTTCCAGACTACTAGTAACGCCTCTGAGGAGTCA-----TACAAAGTGGATGAC	225
Qy	102	AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu	121
Db	226	TTTGCAGCAAAAGTCCTTACAGGAACAACTAACTACAGTTTTTCTCCCACTGTCTAATGC	285
Qy	122	LeuValPheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArg	141
Db	286	ATTGCTTTATCATTTGGTTTGGCAAGCAATGCTATGCCCATCTGGGTCCTTTTTTTTCAG	345
Qy	142	ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu	160
Db	346	ACAAAGAAGAAACATCCTGCTGTGATTATATGTTAACTTGGCATTTGGCAGACCTTCTC	405
Qy	161	PheCysValThrLeuProPheIysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPhe	180
Db	406	TTGCTTGTCTGGTTCCCACTGAAGATTGCATATCATTTAAATGGCAATAACTGGCTGTTT	465
Qy	181	GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle	200
Db	466	GGGGAAGGTCTCTGCAAGAGCTGTTGTTGGATTTTTTATGGAATATGACTGCTCCATT	525
Qy	201	LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr	220
Db	526	CTTTTATGACATGTCTCAGTGTGCAAGGATTGGGNTGTAGTGAACCCCATAGTGCAC	585
Qy	221	ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrVal	240
Db	586	---TCAAGAAAGTCTGAAATTGCCCTGGGCAATCTNCCCTTGTCTATCTGGATCTGATT	642
Qy	241	PheLeuTyrMetLeuProPhePheIleLysGlnGluTyrTyrLeuValGlnPro-As	260
Db	643	TTGGTGGGAACCATTCGGCTGTATCTTGTCAATCANACNGNGTATATTTCAGACCTTTAA	702
Qy	260	pIleThrThr-CysHisAspVal 267	
Db	703	CATCACTACCTGCCATGATGTG 725	

RESULT 32
 AJ447835
 LOCUS
 DEFINITION
 AJ447835
 ACCESSION
 AJ447835.1 GI:20215056
 linear EST 19-APR-2000
 mRNA clone 17n1ir1, mRNA sequence.

```

Qy 240 lPheLeuTyrMetLeuProPhePheLeuLysGlnGluTyrTyrLeuValGlnProAs 260
Db 643 TTTGTTGGAAACCACTCGCTGATCTTGTCAATCAGACNGGATATTTTCAGACCTTAA 702
Qy 260 pileThrThr 263
Db 703 CATCACTACC 712

RESULT 33
BJ068617 633 bp mRNA linear EST 11-DEC-2001
LOCUS BJ068617
DEFINITION laevis cDNA clone XL107d22 5', mRNA sequence.
ACCESSION BJ068617
VERSION BJ068617.1 GI:17494676
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 633)
REFERENCE Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.
EXTRACTED Genes in X. laevis embryo
UNPUBLISHED (2001)
CONTACT: Tadasu Shin-i
NATIONAL INSTITUTE OF GENETICS
1111 Yata, Mishima, Shizuoka 411-8540, Japan
TEL: 81-559-81-6856
FAX: 81-559-81-6855
EMAIL: tshini@genes.nig.ac.jp.
FEATURES
source
1. 633
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL107d22"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 154 a 154 c 143 g 182 t
ORIGIN

Alignment Scores:
Pred. No.: 8.04e-30 Length: 633
Score: 376.00 Matches: 78
Percent Similarity: 55.22% Conservative: 33
Best Local Similarity: 38.81% Mismatches: 65
Query Match: 17.43% Indels: 25
DB: 13 Gaps: 5

US-09-208-629F-6 (1-408) x BJ068617 (1-633)

Qy 33 LeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysPro 52
Db 69 TTATTCCCGACTACAAGCCCGGCTCCAGAGAGAGAGACATGACGGCTGGCGTGAG 128
Qy 53 ThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluLupPheProPhe 72
Db 129 TCTGCTGATCTCGCGCTTT---GCTGCTGCTTCCCAATAGTAGCAGACAGACCTTTC 185
Qy 73 SerAlaLeuGlu-----GlyTrpThrGlyAlaThrIleThrVal 85
Db 186 TACTCCATTGAAAAAGAGAGAGAGTTTTCATGCCCATGGAGTGA----- 230
Qy 86 LysIleLysCysProGluGluSerAlaSerHis-----LeuHis-ValLysAs 101

```

```

Db 231 -----GACGGGCAATGCACCTGGGAAATACTCTCTACACTGTACTCC 272
Qy 101 nAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLe 121
Db 273 CTTTGTTCAAAAGTGTCTCCGATCCGCTGACCCACCGCATTTCTACTCGAGCTACAT 332
Qy 121 uLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheAr 141
Db 333 TATAGTGTTCATCATGCTTACCAAGCAACGCCATTGCACCTGTGGGTCTTTTTCGCG 392
Qy 141 gThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLe 160
Db 393 GACAAAGAGAAGACACCCCTTCCATGATTTATATGGCCCAACCTGGCGTGCAGACTCAT 452
Qy 160 uPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPh 180
Db 453 GTTGTGCATATGGCTCCCATTAAGATTGCATATCATCTGATGGAACCAATGGATTATA 512
Qy 180 eGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSeril 200
Db 513 TGGAGAAGCTTTATGCAAGCTTTTGTATTGGTGTCTTTTATGGGAACATGATTCTCCAT 572
Qy 200 eLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTy 220
Db 573 TCTTTTCATGACATGCCTTAGTGTTCAGAGATACTGGGTGTCATCGTGAACCCCATATCCA 632
Qy 220 r 220
Db 633 C 633

RESULT 34
BI984999
LOCUS BI984999
DEFINITION ful3d01.y3 Campbell zebrafish ovary Danio rerio cDNA clone 5305009
5' similar to SW:PAR2_MOUSE P55086 PROTEINASE ACTIVATED RECEPTOR 2
PRECURSORS ; mRNA sequence.
ACCESSION BI984999
VERSION BI984999.1 GI:16376975
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 591)
REFERENCE Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbratfish@wustl.edu
cDNA Library constructed by Library constructed by Invitrogen and
donated by R. Campbell (Marine Biology Laboratory, Woods Hole, MA).
DNA Sequencing by: Washington University Genome Sequencing Center.
Clone distribution: the I.M.A.G.E. Consortium/LNL, send email to:
info@image.lnl.gov
High quality sequence stop: 441.
FEATURES
source
1. 591
Location/Qualifiers
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5305009"
/clone_lib="Campbell zebrafish ovary"
/sex="female"

```

/dev stage="4-5 months, 1 year and 2 years"
/lab host="DH10B (phage-resistant)"
Note="Organ: ovary (pooled); Vector: pCMV-SPORT6; Site 1:
Site 2: SalI; Whole ovaries collected from zebrafish
aged 4-5 months, 1 year and 2 years. Oligo-dr primed,
directionally cloned. Average insert size 2 kb. Library
constructed by Invitrogen and donated by R. Campbell
(Marine Biology Laboratory, Woods Hole, MA)."

BASE COUNT 129 a 142 c 142 g 178 t

ALIGNMENT SCORES

Pred. No.: 1.35e-29 Length: 591
Score: 373.50 Matches: 75
Percent Similarity: 56.38% Conservative: 31
Best Local Similarity: 39.89% Mismatches: 59
Query Match: 17.32% Indels: 23
DB: 13 Gaps: 5

US-09-208-629F-6 (1-408) x BI984999 (1-591)

Qy 77 GlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHis 96
Db 88 GCCTTTACTGTT-----GAAGAACTAAGGAAGGT 117
Qy 97 LeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIle 116
Db 118 GTATCGGTGACTCCATCGTCTTGGTGATTGACAGCAAGCTCAGCAGATCTTCTC 177
Qy 117 ProAlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrp 136
Db 178 CCCGTGCTCTACATCGTCTTCAGTGTGGCTTGGCCACCACCATGCCATGCCATCTGG 237
Qy 137 MetLeuPhePheArgThr---ArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155
Db 238 GTGTTTCTCTTTAGAACCAAGAACGACCCATCATCCATTTTATGGCCAACTAGCA 297
Qy 156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175
Db 298 CTGGCCGACCTCTCTCTTGTATCTGGATCCCTTGAAATGCTTATCATTTAATGGG 357
Qy 176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
Db 358 AATCATTTGATCTTCGGAGAGCCTTGTGTAAGTGTAGTAGGATTTTCTACGGCAAT 417
Qy 196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
Db 418 ATGTACTGTTCCACAGCTTTTCATCGCATGCATTAGCGTCCAGAGGTACTGGGCCATTGTG 477
Qy 216 HisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGly--- 234
Db 478 CATCCACTTTTCCAGCAG-----AAGAGGAAACACAGGTGGCGGTGTGTGTCA 528
Qy 235 ---LeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db 529 TTGTGCGGTGGTGGTGGTGTG-----TGGGCTTTAAACCGGCGT 567
Qy 254 TyrTyrLeuValGlnProAspIle 261
Db 568 CTGTATCTCTTTACGACCAGACGTG 591

RESULT 35

LOCUS AJ453748 682 bp mRNA linear EST 22-APR-2002
DEFINITION AJ453748 riken1 Gallus gallus cdna clone 3597r1, mRNA sequence.
ACCESSION AJ453748
VERSION AJ453748.1 GI:20263844
KEYWORDS EST.
SOURCE chicken.

ORGANISM Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE

1 (bases 1 to 682)
Buerstedde J.M.
Gallus gallus bursal lymphocyte EST
Unpublished (2002)
COMMENT
Contact: Buerstedde JM
Cellular Immunology
Heinrich-Pette-Institute
Martinistr. 52, 20251 Hamburg, Germany
Email: URL: <http://genetics.hpi.uni-hamburg.de/dt40est.html>.

FEATURES

Location/Qualifiers
1..682
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="3597r1"
/clone_lib="riken1"
/cell_type="bursal lymphocyte"
/dev_stage="2-3 weeks old"
/note="CB inbred strain"

BASE COUNT 160 a 144 c 181 g 197 t
ORIGIN

ALIGNMENT SCORES

Pred. No.: 2.11e-29 Length: 682
Score: 372.50 Matches: 74
Percent Similarity: 55.21% Conservative: 32
Best Local Similarity: 38.54% Mismatches: 81
Query Match: 17.27% Indels: 5
DB: 9 Gaps: 3

US-09-208-629F-6 (1-408) x AJ453748 (1-682)

Qy 62 AlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAla 81
Db 117 GCGCTCCACAGACAGTAGAACAGCAGTTCAAAGCAAGATTTTGTGGCGCAG 176
Qy 82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 177 AAGGTTCCAGATACATAGTAACGCTCTGAGGAGTCA-----TACAAAGTGGATGAC 227
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db 228 TTTGCAGCAAAAGTCCTTACAGGAAACTAACTACAGATTTTCTTCCACTGTGTATGTC 287
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheArg 141
Db 288 ATTGTCTTATCATTTGGTTGGCCAGCAATGCTATGGCCATCTGGGTCTTTTTCAGA 347
Qy 142 ThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeu 160
Db 348 ACAAGAGAAACATCTCTGCTGTGATTATATGTTAACTTGGCATTTGGCAGACCTTCTC 407
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db 408 TTCGTTGCTCGGTCCACCTGAAGATGTCATATCATTTAAATGGCAATACTGGCTGTT 467
Qy 181 GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 468 GGGGAGGCTCTCTGCAAGTGTGTGGATTATTTATGGAATATGTAATGCTGCTCCATT 527
Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db 528 CTTTTATGACATGCTCTCAATGTGCAAGGTATTGGGTTGTAGTGAACCCCATAGTCAC 587
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 588 ---TCAAGAGGAGGTCTGAATTTGGCCCATCTCCCTTGGCTATCTGGATACTGATT 644
Qy 241 PheLeuTyrMetLeuProPhePheIleLeuLysGln 252
Db 645 TTGTTGGGAACCAATTCGCTGTATCTTGTCAATCAG 680

RESULT 36

BE005953

LOCUS BE005953 668 bp mRNA linear EST 05-JUN-2000
 DEFINITION RCO-BN0121-210300-031-g04 BN0121 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE005953
 VERSION BE005953.1 GI:8266186
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 668)
 AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsumura, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=et2-RCO-BN0121-210300-031-g04&t3=2000-03-21&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 35
 High quality sequence stop: 667.
 FEATURES
 source Location/Qualifiers
 1..668
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BN0121"
 /dev_stage="Adult"
 /note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 BASE COUNT 137 a 172 c 157 g 202 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,36e-29 Length: 668
 Score: 370.50 Matches: 82
 Percent Similarity: 53.21% Conservative: 34
 Best Local Similarity: 37.61% Mismatches: 80
 Query Match: 17.18% Indels: 22
 DB: 10 Gaps: 4
 US-09-208-629F-6 (1-408) x BE005953 (1-668)
 Qy 61 GlyAlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGly----- 77
 Db 6 GGGGGCGCATCTGCTAGCAGCTCTCTCTGTCAGTGGCACCACCCAGGTGTATGGCAC 65
 Qy 78 -----TrrP-ThrGlyAlaThrIleThrVallylsileLysCysProGluGluSe 93
 Db 66 ATCCACGCTCACTGGAAAGAGTTACAGTTGAAACAGCTCTTCTCTGTGGATGATTTTC 125
 Qy 93 rAlaSerHisLeuHisVallylsAsnAlaThrMetGlyTyrlLeuThrSerSerLeuSerTh 113
 Db 126 TGCATCT-----GTCCTCACTGGAAACACTGACCAC 155

Qy 113 rLysLeuIleProAlaIleTyrlLeuLeuValPheValValGlyValProAlaAsnAlaVa 133
 Db 156 TGTCTTCCATTCCTTACCAATTGCTACCAATTGTTGTTGGTGGTGGTGGTGGTGGTGGT 215
 Qy 133 lThr-LeuTrrPMet-LeuPhePheArgThrArgSerIleCysThrThrValPheTyrr--- 151
 Db 216 GGCCCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 275
 Qy 152 ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrr 171
 Db 276 GCCAATCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTAT 335
 Qy 172 HisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIle 191
 Db 336 CACATACATGGCAACAACTGGATTATATGGGAAGCTCTTTGTAATGTGCTTATTGGCTTT 395
 Qy 192 PheTyGlyAsnMetTyrrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrr 211
 Db 396 TTCTATGGCAACATGTACTTCTTCCATTCTCTTCATGACCTGGCTTCAGTGTGACAGGTAT 455
 Qy 212 LeuAlaIleValHisProPheThrTyrrArgGlyLeuProLysHisThrTyrrAlaLeuVal 231
 Db 456 TGGGTCTATCTGTAACCCCATGGGGCAC--TCCAGGAAGAGGCAACATTCATTCGTC 512
 Qy 232 ThrCysGlyLeuValTrpAlaThrValPheLeuTyrrMetLeuProPhePheIleLeuLys 251
 Db 513 ATCTCCCTGGCAATATGGTGTCTGTCTGTGTACCATCCCTTTGTATGTGCTGAAG 572
 Qy 252 GlnGluTyrrTyrlLeuValGlnProAspIleThrThrCysHisAspVal 267
 Db 573 CAGACCATCTTTCATCTTCCTGCTGCTGCAACATCAGCAGCTGTCATGATGTT 620

RESULT 37

BF426487 531 bp mRNA linear EST 31-MAY-2002
 DEFINITION df70d03.y1 Xenopus laevis unfertilized egg cDNA library Xenopus laevis cDNA clone XENOPUS SOURCE ID:xltneg006b06 5' similar to SW:PAR2_RAT Q63645 PROTEINASE ACTIVATED RECEPTOR 2 PRECURSOR ;, mRNA sequence.
 BF426487 GI:11437917

ACCESSION

VERSION BF426487.1
 KEYWORDS EST.
 SOURCE African clawed frog.

ORGANISM

Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

REFERENCE

1 (bases 1 to 531)

AUTHORS

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.

TITLE

Washu Xenopus EST project, 1999

JOURNAL

Unpublished (1999)

COMMENT

Contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Library constructed by Bruce Blumberg
 Library normalized by Jihwan Song
 DNA Sequencing by: Washington University Genome Sequencing Center
 This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40RP from Gibco
 High quality sequence stop: 296.

FEATURES

source Location/Qualifiers

1..531
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XENOPUS_SOURCE_ID:xltneg006b06"

/clone.lib="Xenopus laevis unfertilized egg cDNA library"
 /tissue_type="unfertilized egg"
 /lab_host="Top-10 F"

/note=vector: pluescript SK-; Site.1: EcoRI; Site.2:
 XhoI; This library has been used successfully to clone a
 number of full-length cDNAs ranging in size from 1.4 to
 4.5 kb. There are less than 0.5% clones with multiple
 inserts. Since each cDNA has an EcoRI site at its 5' end
 and an XhoI site at the 3' end, these clones can be easily
 identified. One should be suspicious of any clone which
 gives 3 or more bands in an EcoRI-XhoI double digest AND
 has an internal XhoI site. We usually do not further
 characterize any such clones unless the cDNA is known to
 give multiple bands in an EcoRI-XhoI digest. Microplate
 status: 500,000 unamplified cDNAs were mass excised
 (pBluescript SK-) in Xli-Blue using ExAssist phage. The
 resulting single-stranded phagmids were used to infect
 Top10F. Clones were picked into freezing medium (per
 liter 15 g tryptone, 10g yeast extract, 5g NaCl, 36 mM
 K2HPO4, 13.2 mM KH2PO4, 1.7 mM Na-citrate, 0.4 mM MgSO4*7
 H2O, 6.8 mM (NH4)2SO4, 4% w/v glycerol) and grown for 24
 hours. Original library construction by Bruce Blumberg
 (Blumberg et al., 1991 Science 253, 194-196; Hawlet et al.
 , 1995, Genes Dev. 9, 2923-2935). Note: This is a Xenopus
 Gene Collection (XGC) library"

BASE COUNT 133 a 121 c 102 g 175 t
 ORIGIN

Alignment Scores:

Pred. No.: 3.1e-29 Length: 531
 Score: 369.50 Matches: 73
 Percent Similarity: 61.54% Conservative: 31
 Best Local Similarity: 43.20% Mismatches: 58
 Query Match: 17.13% Indels: 7
 DB: 12 Gaps: 5

US-09-208-629F-6 (1-408) x BF426487 (1-531)

Qy 128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
 Db 1 TTACTAGCAACGCGCATTCGATTGGGCTCTTTTTCGGGCAAGAGAAACACCCCT 60
 Qy 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
 Db 61 GCCATGATTATATAGCCCACTTGGCACTGGCAGATTAAATGTTTCATATGCTCCCT 120
 Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
 Db 121 TTTAAATTTGAATATCATCTCAATGGAACAACTGGACTTATGAGAACCTTTATGCAA 180
 Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
 Db 181 GTTTTGATTGGGTTCTTCTATGGGAACATGATGCTCCATTCCTTTTCATGACATGCTT 240
 Qy 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
 Db 241 AGTGTCCAGAGACTAGGGTCATCTGTAACCCCATGTCCTCCAC-----ACAAGAAGAAT 294
 Qy 227 Thr---TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
 Db 295 ACCAAACGTGCTATGATTGTTTCAATTCATATATGGGTGGTCAATTATGCGACACTATT 354
 Qy 246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
 Db 355 CCATTGTACCTGATTAAACCAACTCTGTATCTGTCTGACCTTGGTATACCAACTTGGCAT 414
 Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyr---TyrPheIleSer 284
 Db 415 GACGTCTTCCGCTG-----GATTCTGCAACACTTGACCTGATCTACTACTTGGCT 468
 Qy 285 LeuAlaPhePheGlyPheLeuIlePro 293
 Db 469 CTAGCCCATTTGAGTGTATCTTTCTCCA 495

RESULT 38

AF345566
 LOCUS AF345566 1199 bp mRNA linear HTC 13-JUL-2001
 DEFINITION Homo sapiens putative G-protein-coupled receptor FKSG78 (FKSG78)
 mRNA, complete cds.
 ACCESSION AF345566
 VERSION AF345566.1 GI:13517959
 KEYWORDS HTC.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1199)
 AUTHORS Wang, Y.-g. and Gong, L.
 TITLE Identification of FKSG78, a novel gene encoding a putative
 G-protein-coupled receptor
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1199)
 AUTHORS Wang, Y.-g.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-2001) Beijing FENGKESHENG Function Gene
 Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
 Beijing 100050, P.R. China
 FEATURES
 Location/Qualifiers
 1..1199
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="3"
 /tissue_type="heart"
 /dev_stage="fetus"
 1..1199
 /gene="FKSG78"
 307..1116
 /gene="FKSG78"
 /codon_start=1
 /product="putative G-protein-coupled receptor FKSG78"
 /protein_id="AAK29069.1"
 /db_xref="GI:13517960"
 /translation="MTLTPTFRIVHDAGFGPWFKILCRYTSVLVFANNMTSIVELG
 LISIDRYLKVPRFGDSRMYSITFTKVLSCVVMVIMAVLSLPNIITNGOPTSDNLDH
 CSKLSPGLVKWHTAVYNSCLFVAVLVILGICYIAISRIYHKSRQFTSQSRKRRK
 HNSIRVAVVAFPTCLPFLHCLRFPTFSLDLRLDESQKILYKCEITFLSACNV
 CLDPIVYFMCRSFRSLFKSNIRTRSSIRSLQSVRRSEVRIYYDVTDV"

BASE COUNT

334 a 249 c 245 g 371 t
 ORIGIN

Alignment Scores:

Pred. No.: 1.25e-28 Length: 1199
 Score: 368.50 Matches: 89
 Percent Similarity: 50.33% Conservative: 64
 Best Local Similarity: 29.32% Mismatches: 116
 Query Match: 17.08% Indels: 35
 DB: 11 Gaps: 9

US-09-208-629F-6 (1-408) x AF345566 (1-1199)

Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
 Db 133 AAGAACCACCAC-----CTTCACATGAATTTGACACAAATGTCTTCCCGGTGCTT 183
 Qy 120 TyrLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
 Db 184 TAICTCATATATTTTGGCAAGCACTTCTGCTGAATGGTTAGCAGTGTGG---ATCTTC 240
 Qy 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
 Db 241 TTCACATTAGGAATAAAACACAGCTTCATATTCTATCTCAAAAAACATAGTGGTTCGACAG 300
 Qy 159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
 Db 301 CTCATATGACGCTGACATTTCCATTCGAATAGTCATGATGATGACGAGATTGGACCTTGG 360

```

Qy 179 valPheGlyGluValLeuCysArgAlaThrThrVallePheTyrGlyAsnMetTyrCys 198
Db 361 TACTTCAAGTTATTCTCTGAGATACATCTTCAGTTTGTGTTTATGCAAAATGATGTA 420
Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 421 TCCATCGTTCCTTGGCTGATAGCATGATCGTATCTGATCTGAGGTTGGTCAACCAAT 480
Qy 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 481 GGGGACTCTCGATGATCAGCATAACTTCACGAAGGTTTATCTGTTGTTGTTGGTG 540
Qy 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 541 ATCATGGCTGTTTGTCTTTGCCAAACATCATCTCAACAAATGCT-----CAG 588
Qy 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 589 CCA-----ACAGGAGCAATATCCATGACTGCTCAAAACCTTAAAGTCCTTTGGG 639
Qy 279 LeuTyrTyrPheIleSerLeuAlaPhePheGly-----PheLeuIleProPheVal 295
Db 640 GTCAATGGCATACGGCAGTCACCTATGTAACAGCTGCTGTTGTTGGCGTCTGGTG 699
Qy 296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
Db 700 ATTCTGATCGATGTTATCATAGCCATATCCAGGTACATCCACAAATCCAGCAGGCAATTC 759
Qy 316 Leu-----TrpTyrValLysAlaSerLeuLeuLeuValIle 328
Db 760 ATAAGTCAGTCAACGCCGAAACGAAACATCAACAGAGCATCAGGGTTGTTGGCTGTG 819
Qy 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
Db 820 TTTTACCCTGCTTCTACCATATCACTTGTGCGAGATT-----CCTTTACT 867
Qy 349 TyrAsnAsnThrAspGly-----LeuTyrPheIleTyr 359
Db 868 TTTAGTCACCTAGACAGCTTTTAGATGAATCTGCACAAAAATCCATATTAAGTCAAA 927
Qy 360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
Db 928 GAAATTACACTTTCTTCTGCTGCGTGAATGTTGCTGGAATCAATAATTACTTTTC 987
Qy 380 MetSerLysThr 383
Db 988 ATGTGTAGTCA 999

RESULT 39
AL532537
LOCUS
DEFINITION AL532537 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DM006YC06 5
prime, mRNA sequence.
ACCESSION AL532537
VERSION AL532537.1 GI:12796030
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 904)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DM006YC06"

```

```

/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 260 a 220 c 160 g 264 t
ORIGIN
Alignment Scores: 2.55e-28 Length: 904
Pred. No.: 364.00 Matches: 82
Score: 364.00
Percent Similarity: 49.81% Conservative: 49
Best Local Similarity: 31.18% Mismatches: 88
Query Match: 16.88% Indels: 44
DB: 7
US-09-208-629F-6 (1-408) x AL532537 (1-904)
Qy 106 TyrLeuThrSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db 146 TATGCATCATCAGCAGCGCCAGGATAGTATGCTCTGCATTACAGCCTCGTCTTCATC 205
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 206 ATTGGGCTCGTGGAAACTTACTAGCTTGGTCTGCTCATTTCAAAACACGAAAAAATC 265
Qy 146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
Db 266 AACTCTACCACTCTATTCAACAAATTTGGTGATTCTGATATCTTTTACCACCGCT 325
Qy 165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
Db 326 TTCCCTACAGAAATACCTACTATGCAATGGGCTTTGACATGAGAAATCGGAGATGCTTG 385
Qy 185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAla 204
Db 386 TGTAGGATAACTCGCTAGTGTGTTTATCATCAACATATGTCAGGTGTGAATTTATGACC 445
Qy 205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
Db 446 TGCTGATGATTGACCGCTTCATTGCTGTGGTGCACCTCTAGCGCTACAAACAAGATAAAA 505
Qy 225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
Db 506 AGGATTGAACATGCAAAAGCGGTGTCATATTTGCTGGAATCTAGTATTGTCAGACA 565
Qy 245 LeuProPhePheIle-----LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 566 CTCCCACTCTCATCAACCTATGTCAAACAGGAG-----GCTGAAAGG 610
Qy 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---PropheGlnLeuTyr 280
Db 611 ATTACATGATGATGAGTATCCAAACTTTGAAGAAATAAATCTCTTCCCTGGATTCTG--- 667
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 668 -----CTTGGGCGATGTTTTCATAGGATATGTATCCACTTAATATTCATTCATCTGC 721
Qy 301 TyrAlaAlaIleIle----- 305
Db 722 TATCTCAGATCTGCTGCAAACTTTCAGAACTTTCAGAACTGCGCAAAACCAACCACTCAGTGA 781
Qy 306 -----ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTyrPheValLys 320

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 09:48:59 ; Search time 132.162 Seconds
(without alignments)
4582.672 Million cell updates/sec

Title: US-09-208-629F-6

Perfect score: 2157

Sequence: 1 CSMILQISXRLRDGTGVIKM.....AYLTXXNDLRBQGPSQRT 408

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1055720 seqs, 742224136 residues

Total number of hits satisfying chosen parameters: 2111440

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2_1/USPTO.spool/US09208629/runat_24062003_101628_13027/app.query.fasta_1.1166
-DB=Published Applications_NA -QFMT=fastap -SURF=rnpb -WINMATCH=0.1
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -ENDS=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09208629 @CGN 1 1 153 @runat_24062003_101628_13027
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2142	99.3	1830	US-10-225-567A-301	Sequence 301, App
2	586.5	27.2	1451	US-10-225-567A-299	Sequence 299, App
3	523.5	24.3	4895	US-10-225-567A-515	Sequence 515, App
4	523	24.2	3299	US-10-208-408-24	Sequence 24, Appl

5	523	24.2	3592	9	US-10-225-567A-323	Sequence 323, App
6	488	22.6	2588	10	US-09-943-718-3	Sequence 3, Appli
7	486.5	22.6	1080	10	US-09-943-718-5	Sequence 5, Appli
8	465	21.6	1425	9	US-10-094-417-9	Sequence 9, Appli
9	465	21.6	1955	9	US-10-190-469-2	Sequence 2, Appli
10	465	21.6	2137	9	US-09-782-974C-75	Sequence 75, Appl
11	465	21.6	2137	9	US-10-225-567A-529	Sequence 529, App
12	460	21.3	1080	10	US-09-739-151-1	Sequence 1, Appli
13	413.5	19.2	1020	9	US-10-251-385-181	Sequence 181, App
14	410.5	19.0	1020	9	US-10-251-385-31	Sequence 31, Appl
15	410.5	19.0	1020	10	US-09-788-133-1	Sequence 1, Appli
16	410.5	19.0	1086	9	US-10-222-024-1	Sequence 1, Appli
17	410.5	19.0	1086	9	US-10-251-385-77	Sequence 77, Appl
18	410.5	19.0	1086	9	US-10-251-385-205	Sequence 205, App
19	410.5	19.0	1638	9	US-10-225-567A-111	Sequence 111, App
20	410.5	19.0	2070	9	US-10-225-567A-303	Sequence 303, App
21	401.5	18.6	2424	9	US-10-225-567A-218	Sequence 218, App
22	389	18.0	1301	9	US-10-024-494-7	Sequence 7, Appli
23	387	17.9	1163	9	US-10-121-101B-9	Sequence 9, Appli
24	387	17.9	2299	9	US-10-225-567A-224	Sequence 224, App
25	386.5	17.9	1429	9	US-09-077-173A-1	Sequence 1, Appli
26	384	17.8	1118	9	US-10-225-567A-220	Sequence 220, App
27	382	17.7	1098	9	US-10-225-567A-331	Sequence 331, App
28	376.5	17.5	1041	9	US-09-828-478-1	Sequence 1, Appli
29	376.5	17.5	1041	10	US-09-826-791-5	Sequence 5, Appli
30	376.5	17.5	1041	10	US-09-866-230-6	Sequence 1, Appli
31	376.5	17.5	1260	9	US-09-779-679-1	Sequence 1, Appli
32	376.5	17.5	1260	9	US-09-779-679-24	Sequence 24, Appl
33	376.5	17.5	1430	9	US-09-828-478-3	Sequence 3, Appli
34	376.5	17.5	1700	10	US-09-728-952-26	Sequence 26, Appl
35	376.5	17.5	2807	9	US-10-225-567A-588	Sequence 588, App
36	374.5	17.4	993	10	US-09-826-791-1	Sequence 1, Appli
37	372.5	17.3	1014	9	US-10-225-567A-546	Sequence 546, App
38	371	17.2	1358	9	US-10-167-192-4	Sequence 4, Appli
39	368.5	17.1	1077	10	US-09-826-508-9	Sequence 9, Appli
40	368.5	17.1	1402	9	US-10-225-567A-417	Sequence 417, App
41	368.5	17.1	1481	9	US-10-071-766-136	Sequence 136, App
42	366	17.0	2025	9	US-10-225-567A-216	Sequence 216, App
43	365.5	16.9	1146	9	US-10-225-567A-413	Sequence 413, App
44	362	16.8	1697	9	US-10-109-533A-1	Sequence 1, Appli
45	361	16.7	1444	9	US-09-974-298-85	Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-225-567A-301
; Sequence 301, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn C.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR.
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 301
; LENGTH: 1830
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-301

Alignment Scores:
Pred. No.: 5,896-189
Score: 2142.00
Percent Similarity: 99.26%
Best Local Similarity: 99.26%
Matches: 1830
Conservative: 405
Mismatch: 3

Query Match: 99.30% Indels: 0
DB: 9 Gaps: 0
US-09-208-629f-6 (1-408) x US-10-225-567A-301 (1-1830)
QY 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
DB 88 TGCTCCATGATTTCACAGATTTCATAACGTTTAAAGAGACGGGACTCAGGTTCATCAAAATG 147
QY 21 LysAlaIleuIlePheAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSer 40
DB 148 AAAGCCCTCATCTTTCAGCTGCTGCGCTCCTGCTTCTGTTGCCCATTTTGTCCAGAGT 207
QY 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
DB 208 GGCATGGAAAATGATACAAACACTTGGCAAGCCAACTTACCCATTAAAGACCTTTCGT 267
QY 61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
DB 268 GGAGCTCCCCAAATCTTTTGAAGAGTTCCCTTTTCTGCTTGGAGGCTGGACAGCA 327
QY 81 AlaThrIleThrValLysLysCysProGluSerAlaSerHisLeuHisValLys 100
DB 328 GCCACGATTACTGTAATAATTAAGTCCCTGGAAGAAAGTCTTCACATCTCCATGTGAAA 387
QY 101 AsnAlaThrMetGlyThrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
DB 388 AATGCTACCATGGGTACCTGACACGCTCTTAAGTACTAACTGATCTGCTGCATCTAC 447
QY 121 LeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
DB 448 CTCCTGGTGTCTAGTGTGTGTCGGCCCAATGCTGTGACCTGTGGATGCTTTTCTTC 507
QY 141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaPheLeu 160
DB 508 AGGACGAGATCCATCTGTACCACTGTATTCTACCAACCTCGGCCATTCGAGATTTCCT 567
QY 161 PheCysValThrLeuProPheLeuIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
DB 568 TTTTGTGTACATGGCTTTTAAAGATAGCTTATCTCATCTCAATGGGAACAACTGGGTATT 627
QY 181 GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
DB 628 GGAGAGTCTGTGCGGGCCACACAGTCATCTTCTATGGCAACATGACTGCTCCATT 687
QY 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
DB 688 CTGCTCTTGGCTGTCATCAGCATCAACCGCTACCTGGCCATCTGCTTCATCTTTCACCTAC 747
QY 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
DB 748 CGGGGCTGCCCAAGACACATATGCTTGGTAACATGTGATGATGCTGTGGGCAACAGTT 807
QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
DB 808 TTCTTATATATGCTGCCATTTTTCATCTACAGCAGGAATATTATCTTGTTCAGCCAGAC 867
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
DB 868 ATCACCACCTGCGCATGTATGTTCAACACACTTGGAGTCTCTCATCTCCCTTCCAACTCAT 927
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
DB 928 TACTTCATCTCTTGGCATCTTGTGGATTCTTAATTCATTTGCTTATCATCTACTGTC 987
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
DB 988 TATGACCCCATCATCCGACACTTAATGCATACGATAGATAGTGTGGTATGTTAAG 1047
QY 321 AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
DB 1048 GCGAGTCTCTCTCTCTGTGATTTTACCAATTTGCTTTGCTCCACCAATATTATTTCTT 1107

QY 341 IleIleHisHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
DB 1108 ATTATTCACCATGCTAACTACTACTACAAACACTGATGCTTATATTTATATATCTC 1167
QY 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
DB 1168 ATAGCTTTGTGCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTCTCATG 1227
QY 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
DB 1228 TCRAAAACAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAATGAATCTTAGANA 1287
QY 401 GlnGlyGlnProSerGlnArgThr 408
DB 1288 CAAGGACAGCCATCAGAGAGG 1311
RESULT 2
US-10-225-567A-299
; Sequence 299, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 299
; LENGTH: 1451
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-299
Alignment Scores:
Pred. No.: 5,93e-45 Length: 1451
Score: 586.50 Matches: 141
Percent Similarity: 50.61% Conservative: 67
Best Local Similarity: 34.31% Mismatches: 146
Query Match: 27.15% Indels: 57
DB: 9 Gaps: 14
US-09-208-629f-6 (1-408) x US-10-225-567A-299 (1-1451)
QY 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaLa 27
DB 139 TCACGAGGATGCGGAGCGGCGGCTGCTGGGGCGGCCCATCTGCTAGCA 198
QY 28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
DB 199 GCCTCTCTC-----TCCTGC-----AGTGGCACCATCCAGGAACCAAT 237
QY 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
DB 238 AGATCCTCTAAAGAGAGAGCCCTTATTGGTAAGGTTGATGGCACA-----282
QY 68 GluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrValLysIle 87
DB 283 -----TCCACGTCACCTGA-----AAAGGATTACAGTTGAAACAGTCTTT 324
QY 88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
DB 325 TCTGTGGATGAGTCTTCTGCACTCT-----GTCTTC 354
QY 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
DB 355 ACTGGAAAACACTGACCACGGCTCTCTCCATTTGCTACACAATGTTGTTGGTGGGT 414

```
Qy 128 valproAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db 415 TTGCCAAGTAACGGCATGGCCCTGTGGTCTTTCTTTTCGAACTAAGAAAGACCCCT 474
Qy 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 475 GCTGTGATTACATGGCCAATCTGGCCTTGGTGACCTCTCTCTGTCTGTCTGTGGTCCCT 534
Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 535 TTGAAGATTCCCTATACATACATATGCCAACCACTGGATTATGGGAAGCTCTTTGTAAT 594
Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 595 GTGCTTATTGGCTTTTCTATAGCAACATGACTTCTTCCATTTCTTCTATGACCTGCTC 654
Qy 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db 655 AGTGTGAGAGGATTGGGTATCTGTAACCCCATGGGGCAC---TCCAGGAAGAGGCA 711
Qy 227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db 712 AACATTGCCATTGGCATCTCCCTGGCAATATGCTGTGATTCTGTGGTCAACATCCCT 771
Qy 247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db 772 TTGTATGTCGTGAGCAGACCATCTTCATTCTCTGCCCTGAACATCAGACCTGTCTATGAT 831
Qy 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr----- 280
Db 832 GTT-----TTGCCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 870
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyrCys 300
Db 871 TACTTCTCTCTGGCCATTGGGGCTTTCTGTTCACACCTTCTCTCAGACCTCTGCC 930
Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 931 TATGTGTGTATGATCAGAAATGCTGCGATCTTCTGCCATGGATGAAACTCAGAGAAGAA 990
Qy 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
Db 991 AGGAAGAGGCCATCAAACTATTGTCTGCTGGCCATGACCTGATCTGCTCACT 1050
Qy 335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 1051 CCTAGTAACCTTCTGCTTGTGGTGCAT-----TATTTCTGTATTAGAGCCAGGC 1101
Qy 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1102 CAGAGCCATGCTATGCTGCTGTACATTTAGCCCTCTGCTCTCTACCTTAACAGCTGC 1161
Qy 372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db 1162 ATGACACCCCTTGTCTTATTACTTTGTTTCATGATTTTCAGGGATCATGCAAGAAGCT 1221
Qy 390 TyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1222 CTCCTTTGCCGAGTGTCCGACATGTAAAGCAG 1254
```

RESULT 3

```
US-10-225-567A-515
; Sequence 515, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
```

```
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 515
; LENGTH: 4895
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-515

Alignment Scores:
Pred. No.: 2,26e-38 Length: 4895
Score: 523.50 Matches: 121
Percent Similarity: 53.57% Conservative: 59
Best Local Similarity: 36.01% Mismatches: 145
Query Match: 24.27% Indels: 12
DB: 9 Gaps: 5

US-09-208-629F-6 (1-408) x US-10-225-567A-515 (1-4895)
```

```
Qy 49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
Db 225 CTGGCGGCACCCAGACCCCGCGCTACACGAGAGCGGAGCACCGGA-GGTGGTGAT 283
Qy 68 GluGluPhePro-----PheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrVal 85
Db 284 GACAGCAGCCCTCAATCTGCTGCCCGCGGCTACCCAGGCCAAGTC----- 334
Qy 86 LysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db 335 -----TGCCCAATGACAGTGACACC---CTGGAGCTCCCGACAGCTCACGGGCA 382
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheVal 125
Db 383 CTGCTTCTGGCTGGTGGCCACAGGCTGGTGGCCGCCCTCTATGGGCTGTCTCTGGTG 442
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 443 GTGGGGCTGCCGCCAATGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTG 502
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db 503 CCTCCACCATGCTGTGTGATGAACCTCGCGACTGTGCTGTCTGTGGCCCTGGGCGCTG 562
Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db 563 CCCCCCGGATGCGCTTACCATCTGCTGGCCAGCGCTGGCCCTTCGGGAGGCGCGCTGC 622
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 623 CGCCTGGCCACGCGCCACTCTATGCTACATGATGGCTCAGTGTCTGTCTGTGGCGGCC 682
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
Db 683 GTACGCTGGATGCTGTACCTGCGCTGGTGCAACCCGCTCGGGCCCGCCCTGTGGTGGC 742
Qy 226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db 743 CGCGCGCTGGCCCTTGGACTCTGCATGGCTGTGGCTCATGGCGCGCCCTGGGACATG 802
Qy 246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db 803 CCGCTGACACTGCAGGGCAGACCTTCGCGCTGGCGCTCCGATCCGCTGTCTGTCTGCCAT 862
Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db 863 GACGCGCTGCCCTGGAGCGCCTCCCACTGGCAA---CCGCGCTTCACTGCTGCTG 919
Qy 286 AlaPhePheGlyPheLeuIleProPheValLeuIleTyrCysTyrAlaAlaIleIle 305
Db 920 GCGCTGTTGGCTGTTCTGCTGCCCTGTGGCCATGCTGTGTGTACGGGGCCACCTG 979
Qy 306 ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIle 325
```

Db 980 CACACGCTGGCGCCAGCGCGCTAGCGCCACGCGTGGAGTACCGCGAGTGGG 1039
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHisAla 345
Db 1040 CTGGCTCCGCCGCTGCTCTTCTGTCGCCAGCAACCTGCTGCTGCTGCTACTCG 1099
Qy 346 AsnTyTyTyAsnAsnThrAspGlyLeuTyTyPheIleTyLeuIleAlaLeuCysLeu 365
Db 1100 GACCCAGCCCCAGCGCTGGGGCAACCTCTAGTGGCTTACGTGCCCGCCCTGGCGCTG 1159
Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyTyPheLeuMetSer 381
Db 1160 AGCACCTCAACAGCTGGCGTGGATCCCTTCATCTACTACTACGTGTCG 1207

RESULT 4

US-10-208-408-24
; Sequence 24, Application US/10208408
; Publication No. US20030096272A1
; GENERAL INFORMATION:
; APPLICANT: Schebye, Xiao Min
; TITLE OF INVENTION: GENES REGULATED BY PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR
; FILE REFERENCE: PA-0048-1 US
; CURRENT APPLICATION NUMBER: US/10/208,408
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: 60/308,868
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PERL Program
; SEQ ID NO 24
; LENGTH: 3299
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030096272A1 2456481CB1
US-10-208-408-24

Alignment Scores:
Pred. No.: 1,44e-38 Length: 3299
Score: 523.00 Matches: 129
Percent Similarity: 52.58% Conservative: 75
Best Local Similarity: 33.25% Mismatches: 150
Query Match: 24.25% Indels: 34
DB: 9 Gaps: 12

US-09-208-629F-6 (1-408) x US-10-208-408-24 (1-3299)

Qy 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeu 33
Db 62 GGCGCGGGGGTGGTGGTGGCGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 121
Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThr 53
Db 122 GCCCGCACCGCGCGCGCGGCGGAGCAATCAAAAGCAACAAT-----GCCACC 169
Qy 54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db 170 TTAGATCCCGGTCATTTCTTCTTCAGAACCCCAATGATAATATGAACCATTTTGGGAG 229
Qy 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValIleLys 88
Db 230 GATCAGGAGAAATGAAGTGGTTAACTGAATACAGATTACTCTCCATCAATAAAGC 289
Qy 89 CysProGluSerAlaSerHisLeuHisValIleAsnAlaThrMetGlyTyTyLeuThr 108
Db 290 AGTCCCTTCAAAACAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 349
Qy 109 SerSerLeuSerThrLysLeuIleProAlaIleTyTyLeuLeuValPheValGlyVal 128
Db 350 AGTCTGGGTGACACTTTTTCCTCCCATCTGTGTACACCGGAGTGTGTGTGTGTGTGT 409
Qy 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPheArgThrArgSerIleCys 146

Db 410 CCACATAACATCATGCCCATCGTTGTTCTTCCTGAAATGAAGTCAAGAAG---CG 466
Qy 147 ThrThrValPheTyThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 467 GCGGTGGTGTACATGCTGCACCTGCCAGCGAGATGTCTGTGTGTGTGTGTGTGTGT 526
Qy 167 PheIleAlaTyThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 527 TTTAAGATCAGCATTTACTTTTCCGGCAGTGAATGGCAGTTTGGGTCTGAATGTGTG 586
Qy 187 AlaThrThrValIlePheTyTyGlyAsnMetTyTyCysSerIleLeuLeuLeuAlaCys 206
Db 587 TTCGCTCACTGCAGCATTTTACTGTAAACATGTAGCCCTCTACTCTGCTCATGACAGT 646
Qy 207 SerIleAsnArgTyTyLeuAlaIleValHisPro-----PheThrTyArgGlyLeu 223
Db 647 AGCATTTGACCGGTTTCTGGCTGTGTGTATCCCATGCGAGTCCCTCTCTGGGCTACT 706
Qy 224 ProLysHisThrTyTyAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 243
Db 707 GGAAGGGCTTCCTTC-----ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGG 757
Qy 244 MetLeuProPhePheIleLeuLysGlnTyTyTyTyLeuValGlnProAspIleThrThr 263
Db 758 GTAGTGCCTCTCGTCTCAAGGAGCAACCATCCAGGTGCCCGGCTCAACATCACTACC 817
Qy 264 CysHisAspValHisAsn-----ThrCysGluSerSerSerProPheGlnLeuTyTy 281
Db 818 TGTCTATGATGTCTCAATCAATCAACCTGCTCGAAGGCTACTATCCC-----TACTAC 868
Qy 282 PheIleSerLeuAlaPhePheGlyLeuIleProPheValLeuIleTyTyCysTyTy 301
Db 869 TTTCTAGCTTCTCTGCTGCTCTTCTTTTGGCGCGCATCATTTCCACGGTCTGTAT 928
Qy 302 AlaAlaIleIleArgThrLeuAsn-----AlaTyTyAspHisArg 314
Db 929 GTGCTATCATTCGATGCTTAGCTCTCCGCGAGTTGCCACCGCAGCAAGTCCCG 988
Qy 315 TrpLeuTyTyValLysAlaSerLeuLeuLeuLeuValIlePheThrIleCysPheAla 334
Db 989 GCTTTGTTTC-----CTGTCAGCTGCTGCTTTTCTGTCATCTTTCATTTGCTT 1039
Qy 335 ProSerAsnIleLeuIleIleHisAlaAsnTyTy---TyTyAsnAsnThrAsp 353
Db 1040 CCCAACAAGCTCTCTGATTCGCGCATTTCTCTCTTCTCCTTCTCCTTCTCCTCCT 1099
Qy 354 GlyLeuTyTyPheIleTyTyLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 373
Db 1100 GCTGCTACTTTGCGCTACTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
Qy 374 ProPheLeuTyTyPheLeuMetSer 381
Db 1160 CCCCTAATTACTATTACGCTTCC 1183

RESULT 5

US-10-225-567A-323
; Sequence 323, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Joseph P.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 323


```
; LENGTH: 3592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-323

Alignment Scores:
Pred. No.: 1,63e-38 Length: 3592
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.59% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: Gaps: 11

US-09-208-629f-6 (1-408) x US-10-225-567A-323 (1-3592)

QY 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
Db 348 GGCCCGCGGGCTGCTGCTGGTGGCGCTTTCAGTCTGGCGCGCTGTGTCT 407
QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThr 53
Db 408 GCCCGCACCGCGCGCGCGCGAGCAATCAAAAGCAACAAAT-----GCCACC 455
QY 54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db 456 TTAGATCCCGCGGTCTTCTTCTCAGGAACCCCAATGATAAATATGAACCAATTTTGGGAG 515
QY 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValIleLys 88
Db 516 GATGAGGAGAAATGAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGC 575
QY 89 CysProGluSerAlaSerHisLeuHisValIleAsnAlaThrMetGlyTyLeuThr 108
Db 576 AGTCTCTTCAAAAACAACCTCTCGCATTCATCAGAGATCCCTCCGATATTGGACC 635
QY 109 SerSerLeuSerThrLysLeuIleProAlaIleTyLeuLeuValPheValGlyVal 128
Db 636 AGCTCTGGGTGACACTTTTGTGCCATCTGTACACCGAGTGTGTGTGTGTGTGTGTGT 695
QY 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
Db 696 CCATAACATCATGCGCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 752
QY 147 ThrThrValPheTyThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 753 CGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 812
QY 167 PheIleAlaTyThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 813 TTTAAGATCAGCTATTACTTTTCCGCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 872
QY 187 AlaThrThrValIlePheTyGlyAsnMetTyCysSerIleLeuLeuAlaCysIle 206
Db 873 TTGCTCAGTCAGCATTTTACTTAATGATGATGATGATGATGATGATGATGATGATGAT 932
QY 207 SerIleAsnArgTyLeuAlaIleValHisPro-----PheThrTyArgGlyLeu 223
Db 933 AGCATTCAGCGGTTTCTGGTGTGGTGTATCCCATGCAGTCCCTCTCTGGCGTACTCTG 992
QY 224 ProLysHisThrTyAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTy 243
Db 993 GGAAGGGCTTCTCTC-----ACTTGTCTGGCCATCTGGGCTTGGCCATCGCAGGG 1043
QY 244 MetLeuProPhePheIleLysGlnGluTyTyLeuValGlnProAspIleThrThr 263
Db 1044 GTAGTCCCTCTCTCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACTACC 1103
QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyTyLeuIle 283
Db 1104 TGTTCATGTGTCTCAATGAACCCCTGCTCGAAGG---TACTATGCCTACTACTCTCA 1160
QY 284 SerLeuAlaPheGlyPheProPheValLeuIleIleTyCysTyAlaAla 303
Db 1161 GCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1220
QY 304 IleIleArgThrLeuAsn-----AlaTyAspHisArgTrpLeu 316
Db 1221 ATCATTCGATGCTTGTAGCTCTCCGAGTTCGCAACCGCAGCAAGAGTCCCGGGCTTTG 1280
QY 317 TrpTyValLysAlaSerLeuLeuIleValIlePheThrIleCysPheAlaProSer 336
Db 1281 TTC-----CTGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1331
QY 337 AsnIleLeuIleIleHisAlaAsnTy---TyTyAsnAsnThrAspGlyLeu 355
Db 1332 AACGTCTCTCTGATTCGCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1391
QY 356 TyPheIleTyLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
Db 1392 TACTTTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1451
QY 376 LeuTyPheLeuMetSer 381
Db 1452 ATTACTATTACGCTTC 1469

RESULT 6
US-09-943-718-3
; Sequence 3, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Huffine, Constance P.
; Rossi, Devora L.
; Capone, Myriam
; Hedrick, Joseph A.
; Vicari, Alain
; Gorman, Daniel M.
; Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/009,817
; FILING DATE: 20-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0588K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2589 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1080
```

```

;
; NAME/KEY: misc_feature
; LOCATION: 158
; OTHER INFORMATION: /note= "residues 158, 159, and 276
; probably absent, changing reading frame between those positions;
; sequences provided in SEQ ID NO: 5 and 6"
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-943-718-3

Alignment Scores:
Pred. No.: 1.8e-35 Length: 2588
Score: 488.00 Matches: 117
Percent Similarity: 51.45% Conservative: 60
Best Local Similarity: 34.01% Mismatches: 134
Query Match: 22.62% Indels: 34
DB: 10 Gaps: 4

US-09-208-629F-6 (1-408) x US-09-943-718-3 (1-2588)

QY 56 IleLysThrPheArgGlyAlaProProAsnSerPheGluLupPheProPheSerAlaLeu 75
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1135 GTAAAGATACACAGGAGGCCAT-----1158

QY 76 GluGlyTrpThrAlaThrIleThrValLysLysCysProGluGluSer-----93
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1159 GAAGGCCCTCTGGGTCCC-----ACAGTACAACTCAAGGACGCGAAGTCTCAGACAAG 1212

QY 94 -----AlaSerHisLeuHisVal 99
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1213 CCTAATCCACGAGGCTACCGGGCAAAATCTGTGCCAACGACAGTACACGCTGGAGCTC 1272

QY 100 LysAsnAlaThrMetGlyTrpLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1273 CCGCCAGCTCTCAAGCAGCTGTGGGTGGTCCCTGCCAATGGTGGCGTGTGGTGGTGGCC 1332

QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1333 TATGGGCTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 1392

QY 140 PheArgThrArgSerIleCysThrThrValPheThrThr-AsnLeuAlaIleAlaAspH 159
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1393 ACAAGGTGCCAGCTGCCATCCACCATCTGCTCATGGAACCTGGCAGTGGCTGATCT 1452

QY 159 eLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVa 179
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1453 GCTGTGGCCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGC 1512

QY 179 lPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSe 199
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1513 ATTTGTGAGGCTGCTGGCGGTGGCCACAGCTGCCCTCTATGGCCACATGATGGTTC 1572

QY 199 rIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheTh 219
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1573 AGTGTGCTGCTGGCTGCAGTCACTGGACAGTACCTGGCCCTGGTGGTGGTGGTGGC 1632

QY 219 rTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaTh 239
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1633 GGCCGTGCACTGGGTGCAAGCCCTCACTACTGGACTCTGTTGGTGGCTGGCTCTC 1692

QY 239 rValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPr 259
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1693 TGCAGCCACCTGGCTGCTCTCACTCTGCATCGCAGAACTCCGATTACTGGCTCC 1752

QY 259 oAspIleThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLe 279
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1753 GAT-CGCATGCTGTGTCATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1811

QY 279 uTyrTyrPheIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIle 299
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1812 G----GCCTTCATCTGGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGGCC 1868

QY 299 yrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTrv 319
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 1869 TGTGCTATGGAACCAACCCCTTGTGTCATTGGGGCCAAATGGCCAGCGCTACAGCCATGCAC 1928
QY 319 allYsAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleI 339
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1929 TCAGACTACAGCCCTGTGTACTTGTCTCGGAGTGGCTTCTTACACCTAGCAATGTC 1988
QY 339 leLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleT 359
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1989 TCTGTGCTGCATCTATTCAAACCCGAGCCCTGGGGCAATCTCTATGAGCCT 2048
QY 359 yrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheL 379
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2049 ATGTCCAGCGCTGGCACTCAGCACCTCAACAGCTGGCTAGACCCCTTCTACTACT 2108
QY 379 euMetSer 381
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2109 ATGTGTCC 2116

RESULT 7
US-09-943-718-5
; Sequence 5, Application US/09943718
; Patent No. US20020103361A1
; GENERAL INFORMATION:
; APPLICANT: Huffine, Constance F.
;             Rossi, Devora L.
;             Capone, Myriam
;             Hedrick, Joseph A.
;             Vicari, Alain
;             Gorman, Daniel M.
;             Zlotnik, Albert
; TITLE OF INVENTION: Mammalian Chemokines; Receptors;
;                     Reagents; Uses
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/943,718
; FILING DATE: 30-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/009,817
; FILING DATE: 20-APR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: DX0588K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1077
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-943-718-5

Alignment Scores:

```

Pred. No.: 7,23e-36 Length: 1080
Score: 486.50 Matches: 111
Percent Similarity: 52.96% Conservatives: 59
Best Local Similarity: 34.58% Mismatches: 135
Query Match: 22.55% Indels: 17
DB: 10 Gaps: 2

US-09-208-629F-6 (1-408) x US-09-943-718-5 (1-1080)

```
QY 76 GluGlyTrpThrGlyAlaThrIleThrValLys-IleLysCysProGluSerAlase 95
Db 3 GAAGGCCCTCTGGTCCACAGTACAACTCAAGAGCCGAAGTCTCAGACAGCTAAT 62
QY 95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys-- 114
Db 63 CCACGAGGTACCCGGGCAATCTGTGCCAACAGACAGTGCAGCTGGAGCTCCCGGCC 122
QY 115 -----LeuIleProAlaIleTyrLeuLe 122
Db 123 AGTCTCAAGCACTGCTGGGGTGGTCCCCACAGCTGGTACCTGCCCTCTATGGCT 182
QY 122 uValPheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPheArgTh 142
Db 183 TGTGGTGGCTGTGGGCTGCTGCCAATGGCTGGCGCTGTGGTGTGCCCAAGGCT 242
QY 142 rArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCy 162
Db 243 GCCACGCGCTGCCATCCACCATTCGTCTCATGAACCTGGCAGTGGCTGATCTGTGGC 302
QY 162 sValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 182
Db 303 CTTGGTGTGCCACCACTGCTTACCACCTTCCGCTGGCCAGCGCTGGCCATTTGGTGA 362
QY 182 uValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLe 202
Db 363 GGCTGGCTGGGGTGGCCACACCTGCTCTATGGCCACATGTATGTTTCAGTGTGCT 422
QY 202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgG 222
Db 423 GCTGGCTGCAGTCAGCTTGGACAGATACCTGGCCCTGCTGTCATCTTTGGGCGCCGTG 482
QY 222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
Db 483 ACTGGGTGGTCAACGGCTCACTACTGGACTCTCTTTGGTGGCGCTGCTCTCCAGCCAC 542
QY 242 uTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleTh 262
Db 543 CTGGGCTTGCCTCTCACTCTGCATCGCAGAACTTCCGATTACTGGCTCCGAT-CGCAT 601
QY 262 rThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTrpH 282
Db 602 GCTGTGTATGATGCGCTGCGCTGGCTGAGCAGAACTCCACTGGAGAACG--GCCIT 658
QY 282 eIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleTyrCysTyrA 302
Db 659 CATCTGCTGGCTGTCTCTGGGCTGCTTCTTGGCTGCTGCTGGCCAGCTGTGCTATG 718
QY 302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTyrTrpValLysAlas 322
Db 719 GAACCACTTCTGTGATGGCGGCAATGGCCAGCGCTACACCATGACCTCAGACTGA 778
QY 322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 342
Db 779 CAGCCCTGTGTGTTCTCGGCAGTGGCTTCTTTCACACCTAGCAATGTGCTGCTGGTG 838
QY 342 leHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleA 362
Db 839 TGCACTATTCAAAACCCGAGCCCTGAGGCGCTGGGGCAATCTCTATGGAGCCTATGTGCCA 898
QY 362 laLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 899 GCCTGGCACTACACACCTCAACAGCTGCGTAGACCCCTTTCATCTACTACTATGTGTC 957
```

RESULT 8

US-10-094-417-9
; Sequence 9, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.

; TITLE OF INVENTION: No. US20030045685A1el Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
; FEATURE:
; LOCATION: (241) (1320)
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341

US-10-094-417-9

Alignment Scores:

Pred. No.: 1,05e-33 Length: 1425
Score: 465.00 Matches: 104
Percent Similarity: 51.62% Conservatives: 55
Best Local Similarity: 33.77% Mismatches: 131
Query Match: 21.56% Indels: 18
DB: 9 Gaps: 8

US-09-208-629F-6 (1-408) x US-10-094-417-9 (1-1425)

```
QY 90 ProGluGluSerAlaSerHisLeuHisVal-----LysAsnAlaThr 103
Db 220 CCGGGCGAGACCCCTCCAGGATCGAGTCCCGAACAGCACCGCCCGGACACGCGACG 279
QY 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuVal 123
Db 280 CTGCAGATCTCGGAACCCCGCGATCGCGGTGGCCCTCGCGGTGTACTCGCTGGTG 339
QY 124 PheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPheArg----- 141
Db 340 GCGGCGGTGAGATCCCGGGCAACCTTCTCTGTGGGTGCTGTGCGCGGCATGGGG 399
QY 142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 161
Db 400 CCCAGATCC--CCGTGCGTTCATCTTCATGATCACTGAGCGTCACGACCTGTATGCTG 456
QY 162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db 457 GCCAGCGTGTGCTTCCAAATCTACTACCATTCGACCCGCCACCTCGGTGTTTCGGG 516
QY 182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db 517 GTGCTGCTTTGCAACGTGGTGACCGTGGCTTTTACGCAACATGATTTCCAGCATCTTC 576
QY 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db 577 ACCATGACCTGTATCAGCGTGGAGCGCTTCTCTGGGGTCTGTGTACCGCTCAGCTCCAAG 636
QY 222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db 637 CGTGGCGCGCGCTGTTTACGCGGTGGCGGTGTGTCAGGAGACCTGTGCTGCTCTGCT 696
```

QY 242 LeuTyrMetLeuProPhePheLeuLeuValGlnGluTyrTyrLeuValGlnProAspIle 261
Db : : : : :
QY 697 ACCGCTGTCCCGTGGCGCCAGCATCTACCTACCGGTGACGCCCTGGGCATC 756
Db : : : : :
QY 262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db : : : : :
QY 757 ATCACTGTCTTCAGCGTCTCAAGTGGAGCATGCTCCCGAGCGTGGCCATGTGGGCGTG 816
Db : : : : :
QY 281 TyrPheLeuSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuTyrCys 300
Db : : : : :
QY 817 TTCTCTTTCACCATCTTCCTGCTGCTCTCTCATCCCGTTCGTGATCACCGTGGCTGT 876
Db : : : : :
QY 301 TyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp---HisArgTyr 315
Db : : : : :
QY 877 TACAGGCCACCATCTCAAGCTGTGGCAGCAGGAGCGCGCGCGGAGCAGCGG 936
Db : : : : :
QY 316 LeuTyrTyrValLysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaPro 335
Db : : : : :
QY 937 AGCGCGCGTGGCGCTGGCGCGGTGGTCTTCTGCGCTTTGTACCTGCTTCGCCCC 996
Db : : : : :
QY 336 SerAsnIleLeuLeuLeuLeuHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db : : : : :
QY 997 AACAACTTCGTCTCTCGCGGCACATCGTGAGCGCTGTCTACGGCAAGAGC----- 1050
Db : : : : :
QY 355 LeuTyrPheIleTyrLeuLeuAlaLeuCysGlySerLeuAsnSerCysLeuAspPro 374
Db : : : : :
QY 1051 TACTACAGCTGTACAGCTACAGCTGTGTCTCAGTGTCTCAACAACTGTCTGGACCG 1110
Db : : : : :
QY 375 PheLeuTyrPheLeuMetSerLys 382
Db : : : : :
QY 1111 TTGTTTATTACTTTGGCTCCCGG 1134
Db : : : : :

RESULT 9

US-10-190-469-2
; Sequence 2, Application US/10190469
; Publication No. US20030013155A1
; GENERAL INFORMATION:
; APPLICANT: Gluckmann, Maria A.
; TITLE OF INVENTION: 1400 Receptor, A No. US20030013155A1el G-Protein Coupled Receptor
; FILE REFERENCE: 5800-7, 035800/169196
; CURRENT APPLICATION NUMBER: US/10/190, 469
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US/09/137, 063A
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1955
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1955)
; OTHER INFORMATION: n = a, t, c or g
US-10-190-469-2

Alignment Scores:

Pred. No.:	1.64e-33	Length:	1955
Score:	465.00	Matches:	104
Percent Similarity:	51.62%	Conservative:	55
Best Local Similarity:	33.77%	Mismatches:	131
Query Match:	21.56%	Indels:	18
DB:	9	Gaps:	8

US-09-208-629f-6 (1-408) x US-10-190-469-2 (1-1955)

QY 90 ProGluSerAlaSerHisLeuHisVal-----LysAsnAlaThr 103
Db : : : : :
QY 267 CCGGGGAGACCCCTCCAGGATGCAGTCCCAAGACACCGCGCGGACACGGACG 326
Db : : : : :
QY 104 MetGlyTyrThrSerLeuSerThrLysLeuLeuProAlaIleTyrLeuVal 123
Db : : : : :

Db 327 CTGCAGATGCTCGGGAACCCGCGGATCGCGGTGGCCCTCGCCGTGTGTACTCGTGTG 386
QY 124 PheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArg----- 141
Db : : : : :
QY 387 GCGGCGGTGAGCATCCCGGCAACCTCTTCTCTGTGGTGTGTGTGCGCGCATATGGG 446
Db : : : : :
QY 142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db : : : : :
QY 447 CCCAGATCC---CCGTGCGTCATCTTCATGATCAACCTGAGCGTCACGACCTGATGCTG 503
Db : : : : :
QY 162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db : : : : :
QY 504 GCAGCGGTGTCCTTCCAAATCTACTACCATTCGCAACCCGACCATCGGTATTCGGG 563
Db : : : : :
QY 182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db : : : : :
QY 564 GTGCTGCTTTGCAACGTGTGACCGTGGCTTTTACGCAACATGATTTCCAGCATCTC 623
Db : : : : :
QY 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db : : : : :
QY 624 ACCATGACCTGTATCAGCGTGGAGCGCTTCCTGGGGTCTGTACCGCTCAGCTCCAAG 683
Db : : : : :
QY 222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db : : : : :
QY 684 CGCTGGCGCGCGTCTGTACCGGTGGCGGTGTGCAGGACCTGGCTGTGCTCCTG 743
Db : : : : :
QY 242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db : : : : :
QY 744 ACCGCGCTGTCCCGTGGCGCACCATCTCCTCAAGTGGAGAGTCTCCCGAGCGTGGCGCGG 803
Db : : : : :
QY 262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db : : : : :
QY 804 ATCACTGCTTCAGCGTCTCAAGTGGAGAGTCTCCCGAGCGTGGCGCATGTGGCGCGG 863
Db : : : : :
QY 281 TyrPheLeuSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuTyrCys 300
Db : : : : :
QY 864 TTCTCTTTCACCATCTTCATCTGCTGTCTCTCTCATCCCGTTCGTGATCACCGTGTGT 923
Db : : : : :
QY 301 TyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp---HisArgTyr 315
Db : : : : :
QY 924 TACAGCGGCACCATCTCAAGCTGTGGCAGCAGGAGCGCGCGCGGAGCAGCGG 983
Db : : : : :
QY 316 LeuTyrTyrValLysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaPro 335
Db : : : : :
QY 984 AGCGCGCGTGGCGCTGGCGCGGTGTCTGCTGGCTTTGTACCTGCTTCGCCCC 1043
Db : : : : :
QY 336 SerAsnIleLeuLeuLeuHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db : : : : :
QY 1044 AACAACTTCGTCTCTCGCGCACATCGTGAGCGCGCTGTCTACGGCAAGAGC----- 1097
Db : : : : :
QY 355 LeuTyrPheIleTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db : : : : :
QY 1098 TACTACCGCTGTACAGCTCAGCTGTGTCTCAGTGTCTCAACAACTGTCTGGACCGG 1157
Db : : : : :
QY 375 PheLeuTyrPheLeuMetSerLys 382
Db : : : : :
QY 1158 TTGTTTATTACTTTGGCTCCCGG 1181
Db : : : : :

RESULT 10

US-09-782-974C-75
; Sequence 75, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogel, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1el G Protein Coupled Receptor
; FILE REFERENCE: 41USPHERM311
; CURRENT APPLICATION NUMBER: US/09/782, 974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165, 838
; PRIOR FILING DATE: 1999-11-16

;; PRIOR APPLICATION NUMBER: 09/714,449
;; PRIOR FILING DATE: 2000-11-16
;; PRIOR APPLICATION NUMBER: 60/198,568
;; PRIOR FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: 60/166,071
;; PRIOR FILING DATE: 1999-11-17
;; PRIOR APPLICATION NUMBER: 60/166,678
;; PRIOR FILING DATE: 1999-11-19
;; PRIOR APPLICATION NUMBER: 60/173,396
;; PRIOR FILING DATE: 1999-12-28
;; PRIOR APPLICATION NUMBER: 60/184,129
;; PRIOR FILING DATE: 2000-02-22
;; PRIOR APPLICATION NUMBER: 60/185,421
;; PRIOR FILING DATE: 2000-02-28
;; PRIOR APPLICATION NUMBER: 60/185,554
;; PRIOR FILING DATE: 2000-02-28
;; PRIOR APPLICATION NUMBER: 60/186,530
;; PRIOR FILING DATE: 2000-03-02
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 192
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 75
;; LENGTH: 2137
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-782-974C-75

Alignment Scores:
Pred. No.: 1.86e-33 Length: 2137
Score: 465.00 Matches: 104
Percent Similarity: 51.62% Conservative: 55
Best Local Similarity: 33.77% Mismatches: 131
Query Match: 21.56% Indels: 18
DB: Gaps: 8

US-09-208-629F-6 (1-408) x US-09-782-974C-75 (1-2137)

QY 90 ProGluSerAlaSerHisLeuHisVal-----LysAsnAlaThr 103
Db 195 CCGGGGAGGAGCCCTCCAGGATGCGAGTCCCGAAGACGACCGCGCGGACACGCGAGC 254
QY 104 MetGlyTyrLeuThrSerLeuSerThyLysLeuLeuProAlaLeuVal 123
Db 255 CTCGAGATGTCGGGAACCGCGGATCGCGGTGGCCCTGCGGTGTGTACTCGCTGGTG 314
QY 124 PheValValGlyValProAlaAsnAlaValThyLeuTyrMetLeuPheArg----- 141
Db 315 GCGGGGTACGATCCCGGCAACCTCTCTGTGGGTGTGTGCGCGGCATGGGG 374
QY 142 ThrArgSerIleCysThrThrValPheThyThrAsnLeuAlaLeuAlaAspPheLeuPhe 161
Db 375 CCCAGATCC--CCGTGGGTCACTTCATGATCACTGAGCTGACGGACCTGTGCTG 431
QY 162 CysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db 432 GCCAGCGTGTGGCTTTTCAAACTACTACCATTTCAACCGCCACCACTGGGTATTCGGG 491
QY 182 GluValLeuCysArgAlaThrValIlePheThyGlyAsnMetTyrCysSerIleLeu 201
Db 492 GTCTGCTTTGCAACGTGGTGGCGCTTTTACGCAAACTATGATTCACGATCCTC 551
QY 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db 552 ACCATGACCTGTATCAGCGTGGAGCGCTTCTGGGGTCTGTACCGCTCAGCTCCAG 611
QY 222 GlyLeuProIysHisThrTyrAlaLeuValThyCysGlyLeuValTrpAlaThrValPhe 241
Db 612 CGTGGCGCGCGCTGTTACGCGGTGGCGCGGTGTGTCAGGACCTGGCTGCTCCTG 671
QY 242 LeuTyrMetLeuProPheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 672 ACCGCCCTGTCCCGTGGCGGACCGATCTCACCTACCGGTGTCGCGCCCTGGGCATC 731

QY 262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db 732 ATCACCTGCTTCAGCTCCTCAAGTGGACGATCTCCCGACGCTGGCCATGTGGCGGTG 791
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuIleTyrCys 300
Db 792 TTCCTCTTCCATCTTCATCTCCTGCTGCTCCTCATCCGTTCTGTATCACCCTGGCTGT 851
QY 301 TyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp---HisArgTyr 315
Db 852 TACACGGCCACCATCTCAAGCTGTTGCGACGAGGAGGCGACGCGCGGAGCAGCG 911
QY 316 LeuTyrTyrValysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaPro 335
Db 912 AGGCGCGGTGGCGCTGGCGGTGGTCTTGTGCTGCTTGTACCTGCTTGGCGCC 971
QY 336 SerAsnIleIleLeuIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 972 AACAACTTCTGCTCTCTGCGCACATCGTGAGCGGCTGTCTTACGCGCAAGAGC----- 1025
QY 355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db 1026 TACTACACGCTGACAGCTCAGCTGTGTCTCAGCTGCTCAACACTGTCTGGACCG 1085
QY 375 PheLeuTyrPheLeuMetSerLys 382
Db 1086 TTGTTTATTACTTTTGGTCCCGG 1109

RESULT 11

US-10-225-567A-529
; Sequence 529, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christina L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 529
; LENGTH: 2137
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-529

Alignment Scores:
Pred. No.: 1.86e-33 Length: 2137
Score: 465.00 Matches: 104
Percent Similarity: 51.62% Conservative: 55
Best Local Similarity: 33.77% Mismatches: 131
Query Match: 21.56% Indels: 18
DB: Gaps: 8

US-09-208-629F-6 (1-408) x US-10-225-567A-529 (1-2137)

QY 90 ProGluSerAlaSerHisVal-----LysAsnAlaThr 103
Db 195 CCGGGGAGGAGCCCTCCAGGATGCGAGTCCCGAAGACGACCGCGCGGACACGCGAGC 254
QY 104 MetGlyTyrLeuThrSerSerLeuSerThyLysLeuLeuProAlaIleTyrLeuVal 123
Db 255 CTCGAGATGTCGGGAACCGCGGATCGCGGTGGCCCTGCGGTGTGTACTCGCTGGTG 314
QY 124 PheValValGlyValProAlaAsnAlaValThyLeuTyrMetLeuPheArg----- 141
Db 315 GCGGGGTACGATCCCGGCAACCTCTCTCTGTGGGTGTGTGCGCGGCATGGGG 374

```
QY 142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
Db 375 CCCAGATCC---CCGTCGGTCACTTCATGATCAACTGAGCGTCACGGACCTGATGCTG 431

QY 162 CysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
Db 432 GCCAGCGGTGTGCTTCCAAATCTACTACCATTCGACCGCCACCACTGGGTATTTCGGG 491

QY 182 GluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
Db 492 GTGCTGTTTGAACGTGTGACCGGTGCGCTTTTACGAAACATGATTCCAGCATCCCTC 551

QY 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
Db 552 ACCATGACCTGTATCAGCGGAGCGCTTCTCGGGGTCTGTACCGCTCAGCTCCAAG 611

QY 222 GlyLeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
Db 612 CGCTGGCGCGCGCTGCTTACGGCGGTGGCGCGTGTGACGGGACCTGGCTGTCTCCTG 671

QY 242 LeuTyrMetLeuProPhePheIleLeuIlysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 672 ACGCCCTGTCCCGCTGGCGCGACGATCTCACCTACCGGTGACGCGCCCTGGGCGATC 731

QY 262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280
Db 732 ATCACCTGTCTCAGCTCTCAAGTGACGATGCTCCCGACGCTGGCGCTGTGGCGCGT 791

QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 792 TTCTCTTTCACCATCTTCATCTCTGCTGCTTCTCATCTCCGTTCTGTGATCACCCTGGTGT 851

QY 301 TyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
Db 852 TACAGCGCCACCATCTCAAGCTGTTCGCGACGAGGAGCGCGACCGCGCGGAGCAGCGG 911

QY 316 LeuTrpTyrValIysAlaSerLeuLeuIleValIlePheThrIleCysPheAlaPro 335
Db 912 AGCGCGCGGTGGCGCGCTGGCGGTGGTCTTCTGCTGCGCTTGTCACTGCTTCGCGCCC 971

QY 336 SerAsnIleIleLeuIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 972 AACAACTTCGTGCTCTCGCGGCACATCGTGAGCGCGCTGTCTTACGGCAAGAGC----- 1025

QY 355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db 1026 TACTACACGTGTACAACTCAGCTGCTGTGCTCAGCTGCTCAACAACTGTCTGGACCG 1085

QY 375 PheLeuTyrPheLeuMetSerLys 382
Db 1086 TTGTGTTATTACTTTGGTCCCGG 1109
```

RESULT 12

```
US-09-739-151-1
; Sequence 1, Application US/09739151
; Patent No. US20010029032A1
; GENERAL INFORMATION:
; APPLICANT: Yuan Zhu
; APPLICANT: Xiaotong Li
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: PAUL, A G-PROTEIN COUPLED RECEPTOR
; TITLE OF INVENTION: POLYPEPTIDE
; FILE REFERENCE: GP-70567-C1
; CURRENT APPLICATION NUMBER: US/09/739,151
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/413,534
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: 60/103,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1080
```

```
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-739-151-1

Alignment Scores:
Pred. No.: 2,066-33 Length: 1080
Score: 460.00 Matches: 100
Percent Similarity: 52.92% Conservative: 54
Best Local Similarity: 34.36% Mismatches: 125
Query Match: 21.33% Indels: 12
DB: 10 Gaps: 7
```

US-09-208-629F-6 (1-408) x US-09-739-151-1 (1-1080)

```
QY 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyr 120
Db 31 AACGGACGCTCAGATGCTGCGGAACCCGGCGATCGCGGTGGCCCTGCCCGTGGTGATC 90

QY 121 LeuLeuValPheValValGlyValProAlaAlaValThrLeuTrpMetLeuPhePhe 140
Db 91 TCGCTGGTGGCGCGGTGAGCATCCCGGGCAACCTCTTCTCTCTGTGGGTGTGTGCCCG 150

QY 141 Arg-----ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAsp 158
Db 151 CGCATGGGGCCAGATCC---CCGTCGGTCACTTTCATGATCAACCTGAGCGTCACGGAC 207

QY 159 PheLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db 208 CTGATGTGGCGCAGCGTGTGCTTCCAAATCTACTACCATTCGACCGCCACCACTGG 267

QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 268 GTATTGGGGTCTGCTTTGCAACGTGGTGACCGTGGCCTTTTACGCAACATGATTCC 327

QY 199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 328 AGCATCTTCACCATCACTGTATCAGCGTGGAGCGCTTCTCGGGGTCTGTACCGCGCTC 387

QY 219 ThrTyrArgGlyLeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 388 AGCTTCAGCGGTGGCGCGCGTGTGTACGGCGGTGGCGCGGTGTGTGAGGACCTGGCTG 447

QY 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 448 CTGCTCTCCTACCGCTGCTGCTGGCGCGCACCGATCTCACCTTACCGGTGCGACGCC 507

QY 259 ProAspIleThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPhe 277
Db 508 CTGGGCATCATCACTGCTTCGACGCTCTCAAGTGGACGATGCTCCCCAGCGTGGCCATG 567

QY 278 GlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIle 297
Db 568 TGGGGGTGTCTCTTCACTATCTCATCTGCTGTCTCATCCCGTGTGTGATCACC 627

QY 298 IleTyrCysTyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp--- 312
Db 628 GTGGCTGTGTACCGGCCACCATCTCAAGCTGTTCGCGCACGGAGGAGCGCGACGCCCG 687

QY 313 HisArgTrpLeuTrpTyrValIysAlaSerLeuLeuIleLeuValIlePheThrIleCys 332
Db 688 GAGCAGCGGAGGCGCGGTGGCGCTTTCGCGCGGTGTCTTGTGCGCTTGTGTCACTGC 747

QY 333 PheAlaProSerAsnIleIleLeuIleHis---HisAlaAsnTyrTyrTyrAsnAsn 351
Db 748 TTGCGCCCCCAACAACCTTCGTCTCTCGCGCACATCGTGAGCGCGCTGTCTTACGGCAAG 807

QY 352 ThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 808 AGC-----TACTACACCGTGTACAAGCTCACGCTGTGTCTCAGCTGCCTCAACAAC 861

QY 372 LeuAspProPheLeuTyrPheLeuMetSerLys 382
Db 862 CTGGACCGCGTTTGTATTACTTTTGGTCCCGG 894
```

RESULT 13
US-10-251-385-181
; Sequence 181, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 181
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-181

Alignment Scores:
Pred. No.: 3,88e-29 Length: 1020
Score: 413.50 Matches: 108
Percent Similarity: 45.11% Conservative: 49
Best Local Similarity: 31.03% Mismatches: 130
Query Match: 19.17% Indels: 61
DB: 9 Gaps: 13

US-09-208-629F-6 (1-408) x US-10-251-385-181 (1-1020)

QY 62 AlaProAsnSerPheGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 19 GCTCCCGGAGTCTGATCACCACCTTCTCCCTGGCCAGGAG----- 63
QY 82 ThrIleThrValLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 64 -----CAATGGCCAGGAG----- 78
QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeu 121
Db 79 -----ACGCCACTGGAGAACATGCTGTCGCCTCTCTACCTT 117
QY 122 LeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db 118 CTGGATTTATCTGGCTTTAGTTGGCAATACCTGGCTCTGTGG---CTTTTCATCCGA 174
QY 142 ThrArgSerIleCysThr-----ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db 175 GACCACAGTCCGGGACCCCGGCAAGCTGTTCTGATGATCTGGCCGTGGCCGACTTG 234
QY 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpVal 179
Db 235 TCGTGGCTGTGCTGCTGCCCGCCCGCTGCTTACCACCTCTCTGGGAACCACTGGCCA 294
QY 180 PheGlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db 295 TTTGGGAAATCGATCGCTCTCACCGGCTCTCTCTTACCTCAACATGTACGCCAGC 354
QY 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 355 ATCTACTCTCTACCTCATCAGCGGCGGCGGTTCTCTGGCCATTTGACCCGGTCAAG 414
QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 415 TCCTCAAGCTCCGAGCCCTCTAGGCACACCTGGCTGTGCTTCTCTGTGGTGGTG 474
QY 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db 415 TCCTCAAGCTCCGAGCCCTCTAGGCACACCTGGCTGTGCTTCTCTGTGGTGGTG 474

Db 475 GTGGCTGTGGCCATGGCCCGCTGTGGTGAGCCACAG----- 513
QY 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeu 279
Db 514 ACGTGCAGACCAACACACAGGGTG-----GTCTGC-----CTGCAGCTG 552
QY 280 Tyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuLeu 292
Db 553 TACCGGAGAGAGCGCTCCACCACCATGCTGGTGTCTCCCTGGCA---GTGGCCTTACCCCTTC 609
QY 293 PropheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn----- 309
Db 610 CCGTTCATCACCAGGTCTACCTGCTACCTGCTGATCATCCGAGCCTCGGCGAGGCGCTG 669
QY 310 AlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu---LeuIleLeuValIle 328
Db 670 CGTGTGAGAGCGCTCAAGACCAAGGCAACGATGATCGCCATGCTGTGGCATC 729
QY 329 PheThrIleCysPheAlaProSerAsnIle-----IleLeuIleIleHisAla 345
Db 730 TTCTGTGTCTGTCTGTGGCTTACCACGTCAACCGCTCGCTGTACGTGTGCTACTACCGC 789
QY 346 AsnTyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 790 AGCATGGGGCTCTCTGGCCACCCAGCGCATCTGGCCCTGGCAACCGCATCACCTCC 849
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 850 TGCTTCACCAAGCTCAACGGGCACTCGACCCCATCATGTAATTCTTCTGTGGCTGAGAAG 909
QY 384 ArgAsnHisSerThrAlaTyrLeu 391
Db 910 TTCCGCGCAGCGCTGTGCAACTTG 933

RESULT 14

US-10-251-385-31
; Sequence 31, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 31
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-31

Alignment Scores:
Pred. No.: 7,36e-29 Length: 1020
Score: 410.50 Matches: 107
Percent Similarity: 44.54% Conservative: 48
Best Local Similarity: 30.75% Mismatches: 132
Query Match: 19.03% Indels: 61
DB: 9 Gaps: 12

US-09-208-629F-6 (1-408) x US-10-251-385-31 (1-1020)

QY 62 AlaProAsnSerPheGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 19 GCTCCCGGAGTCTGATCACCACCTTCTCCCTGGCCAGGAG----- 63
QY 82 ThrIleThrValLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101

[illegible]

```
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P79011
; CURRENT APPLICATION NUMBER: US/09/788,133
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1020)
; US-09-788-133-1

Alignment Scores:
Pred. No.:      7,36e-29      Length:      1020
Score:          410.50       Matches:    107
Percent Similarity: 44.54%   Conservative: 48
Best Local Similarity: 30.75% Mismatches:   132
Query Match:     19.03%     Indels:      61
DB:              12        Gaps:      12

US-09-208-629F-6 (1-408) x US-09-788-133-1 (1-1020)

Qy      62 AlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
      |||||
Db      19 GCTCCCCCAGGTCTGATCACCAACTTCTCCCTGGCCACGGCAGAG----- 63

Qy      82 ThrIlethrVallYsIlleLysCySProGluGluSerAlaSerHisLeuHisValLysAsn 101
      |||||
Db      64 -----CAATGTGGCCAGGAG----- 78

Qy      102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
      |||||
Db      79 -----ACGCCACTGGAGAACAATGCTGTTGCCTCTCTTCTACCTT 117

Qy      122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
      |||||
Db      118 CTGAATTATTCCTGGCTTTTAGTTGGCAATAACCTCGCTCTGTGG--CTTTTCATCCGA 174

Qy      142 ThrArgSerileCysThr-----ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
      |||||
Db      175 GACCACAAAGTCGGGACCCGGCCAAAGTGTTCTGTATGCAATGGCCGTGGCCACATGG 234

Qy      160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
      |||||
Db      235 TCGTGGCTGTGGTCTGTCCGCCACCCGCTGTGTCTACCACTTCTCTGGGAACCACTGGCCA 294

Qy      180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
      |||||
Db      295 TTGTGGGAAAATCGCATCCGCTCTCACC GGCTTCTCTTCTACCTCAAATGATACGCCACG 354

Qy      200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPropHeThr 219
      |||||
Db      355 ATCTACTTCTCACCTGCATGCAGCCGACCGCTTCTCTGGGCATTTGTGACCCGGTCAAG 414

Qy      220 TyrArgGlyLeuProLysHisThrTyxAlaLeuValThrCysGlyLeuValTrpAlaThr 239
      |||||
Db      415 TCCCTCAAGCTCCGACGGCCCTCTACGCACACCTGGCCTGTGCCTTCTGTGGGTGGTG 474

Qy      240 ValPheLeuThrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
      |||||
Db      475 GTGGCTGTGGCCATGGGCCCGCCCTGCTGTGTGAGCCACAG----- 513

Qy      260 AspileThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
      |||||
Db      514 ACCGTGCAGACCAACACACCGTG-----GTCTGC-----CTGCAGCTG 552

Qy      280 Tyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
      |||||
Db      553 TACCGGGAAGGGCTCCCAACATGCTGTGTGCTCCCTGGCA---GTGGCCTTCACTTTC 609
```


Alignment Scores:
Pred. No.: 8,04e-29 Length: 1086
Score: 410.50 Matches: 99
Percent Similarity: 47.87% Conservatives: 58
Best Local Similarity: 30.18% Mismatches: 114
Query Match: 19.03% Indels: 57
DB: Gaps: 10

US-09-208-629F-6 (1-408) x US-10-251-385-77 (1-1086)

QY 106 TyrLeuThrSerLeuSerThrLysLeuLeuProAlaIleTyrLeuLeuValPheVal 125
DB 70 TATGCACATCACAGCAGCGGATAGTAACTGCTGCATTACAGCCTCGTCTTCATC 129
QY 126 ValGlyValProAlaAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIle 145
DB 130 ATTGGGCTCGTGGGAACTTACTAGGCTTGGTCTGCTCATTTGTTCAAAACAGGAAAAATC 189
QY 146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
DB 190 AACTCTACCACTCTATTCACAAATTTGGTGATTTCTGATATATCTTTTACCAGGCT 249
QY 165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
DB 250 TTGCTTACAGCATAGCTTACTATGCAATGGGCTTTGACTGGAGATCGGAGTGCCTTG 309
QY 185 CysArgAlaThrThrValPheTyrGlyAsnMetTyrCysSerIleLeuLeuAla 204
DB 310 TGTAGGATAACTCGCTAGTGTGTTTATCATCAACATATGCAAGTGTGAACCTTATGACC 369
QY 205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrThrTyrArgGlyLeuPro 224
DB 370 TGCTGAGTATGACCGCTTCTATGCTGGTGGTGCACCTCTACGCTACAAAGATAAAA 429
QY 225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
DB 430 AGGATTGAACATGCAAAAGCGCTGTCATATTTGCTGATTTCTAGTATTTGCTCAGACA 489
QY 245 LeuProPhePheIle-----LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
DB 490 CTCCTCCTCTCTCAACCTATGTCAAAGAGGAG-----GCTGAAGG 534
QY 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer-----PropheGlnLeuTyr 280

QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
DB 592 ---CTTGGGCGATGTTTCATAGGATATGTACTCCACTTATATCATCTCATCTGCG 645
QY 301 TyrAlaAlaIleIle----- 305
DB 646 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACCAAAACCCACTCACTGAGAAA 705
QY 306 -----ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTyrTrpValLys 320
DB 706 TCTGGTGTAACAAAAGGCTCTCAACAAATT----- 738
QY 321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
DB 739 -----ATTCTTATTTGTTGTTGTTCTCTGTTTCAACCTTACCATGTTGCAATT 792
QY 341 IleIleHis-----HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
DB 793 ATTCACATATGATTGAAGAGCTTCTGTTTCTCTAAATTTCTGGAATGTAGCCAAAGACAT 852
QY 355 LeuTyrPheIleTyrLeu-----IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
DB 853 TCGTTCAGATTTCTCTGCACTTTTACAGTATGCTGTATGAACCTTCAATGCTCATGGAC 912
QY 374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393

DB 913 CCTTTTATCTACTCTTTTGCATGTAAA-----CGGTATAAGAGAAAG 954

QY 394 *****AsnAspLeuArgGluGln 401
DB 955 GTTATGAGGATGCTGAAACGGCAA 978

RESULT 18
US-10-251-385-205
; Sequence 205, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 205
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-205

Alignment Scores:
Pred. No.: 8,04e-29 Length: 1086
Score: 410.50 Matches: 102
Percent Similarity: 50.47% Conservatives: 58
Best Local Similarity: 32.18% Mismatches: 122
Query Match: 19.03% Indels: 35
DB: Gaps: 11

US-09-208-629F-6 (1-408) x US-10-251-385-205 (1-1086)

QY 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
DB 70 TATGCACATCACAGCAGCGGATAGTAACTGCTGCATTACAGCCTCGTCTTCATC 129
QY 126 ValGlyValProAlaAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIle 145
DB 130 ATTGGGCTCGTGGGAACTTACTAGCTTGGTGGTCTGCTCATTTGTTCAAAACAGGAAAAATC 189
QY 146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
DB 190 AACTCTACCACTCTATTCACAAATTTGGTGATTTCTGATATATCTTTTACCAGGCT 249
QY 165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
DB 250 TTGCTTACAGCATAGCTTACTATGCAATGGGCTTTGACTGGAGATCGGAGTGCCTTG 309
QY 185 CysArgAlaThrThrValPheTyrGlyAsnMetTyrCysSerIleLeuLeuAla 204
DB 310 TGTAGGATAACTCGCTAGTGTGTTTATCATCAACATATGCAAGTGTGAACCTTATGACC 369
QY 205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrThrTyrArgGlyLeuPro 224
DB 370 TGCTGAGTATGACCGCTTCTATGCTGGTGGTGCACCTCTACGCTACAAAGATAAAA 429
QY 225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
DB 430 AGGATTGAACATGCAAAAGCGCTGTCATATTTGCTGATTTCTAGTATTTGCTCAGACA 489
QY 245 LeuProPhePheIle-----LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
DB 490 CTCCTCCTCTCTCAACCTATGTCAAAGAGGAG-----GCTGAAGG 534
QY 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer-----PropheGlnLeuTyr 280

```
Db 535 ATTACATGCGATGAGTATCCAACTTTGAGAAACATAAATCTCTCCCTGGATTCTG--- 591
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 592 -----CTTGGGCGATGTTTCATAGATATGTAATCCATTTATATCATCTCATCTGC 645
Qy 301 TyrAlaAlaIle-----IleArgThrLeu-----AsnAlaTyrAspHisArg 314
Db 646 TATTCTCAGATGCTGCAAACTCTTCAGAACTGCGCAAAACCAAAACCCACTCACTGAGAAA 705
Qy 315 TrpLeuTrpTyrValIlyAlaSer-----LeuLeuIleValIlePheThrIle 331
Db 706 TCTGGTGTAACAAAAGGCTAAACACAAATATTCTTATTATTGTTGTTGTTCTC 765
Qy 332 CysPheAlaProSerAsnIleLeuIleIleHis-----HisAla 345
Db 766 TGTTTCACACCTTACCATGTTGCAATATTCAACATATGATTAAAGAGCTCTGTTCTCT 825
Qy 346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu---IleAlaLeuCys 364
Db 826 AATTTCCTGGAATGTAGCCAAAGACATTCGTTCCAGATTCTCTGCACITTTACAGTATGC 885
Qy 365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerIlyThrArg 384
Db 886 CTGATGAACCTCAATTGCTGCATGGACCCCTTTTCTACTCTTTCATGTATAA----- 939
Qy 385 AsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGluGln 401
Db 940 -----CGGTATAAGAAAGGTTATGAGGATGCTGAACGGCAA 978
```

RESULT 19

```
US-10-225-567A-111
; Sequence 111, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 111
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-111
```

```
Alignment Scores:
Pred. No.: 1,43e-28 Length: 1638
Score: 410.50 Matches: 99
Percent Similarity: 47.87% Conservative: 58
Best Local Similarity: 30.18% Mismatches: 114
Query Match: 19.03% Indels: 57
DB: Gaps: 10
```

US-09-208-629f-6 (1-408) x US-10-225-567A-111 (1-1638)

```
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db 103 TATGCACATCACAGCGCGGAGATAGTAATGCCCTCTGCATTACAGCCCTCGCTTCATC 162
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 163 ATTGGGCTCGTGGAAACTTACTAGCCCTTGCTGCTCATGTTTCAAAACAGGAAAAATC 222
Qy 146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
```

```
Db 223 AACTCTACACCCCTCTATTCAACAAATTTGGTGATTCTTGATATACTATTTTACCACCGCT 282
Qy 165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
Db 283 TTGCTTACAGTAATACCTTACTATGCAATGGGCTTTGACTGGAGAATCGGAGATGCTTGG 342
Qy 185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAla 204
Db 343 TGTAGGATAACTCGCTAGTGTGTTTACATCAACACATATGCAGGTGTGAACCTTTATGACC 402
Qy 205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
Db 403 TGCTTGAGTATTGACCGCTTCAATGCTGTGGTGACCCCTCTACGCTACAAACAGATAAAA 462
Qy 225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
Db 463 AGGATTGAACATGCCAAAAGCGGTGTCATATTGCTGGATTCTAGTATTGTCAGACA 522
Qy 245 LeuProPhePheIle-----LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 523 CTCCCACTCTCATCAACCCCTATGTCAAAGCAGGAG-----GCTGAAAGG 567
Qy 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer---ProPheGlnLeuTyr 280
Db 568 ATTACATGATGATGATATCCAAACTTTGAAGAAACTAAATCTCTTCCCTGGATTCTG--- 624
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 625 -----CTTGGGCGATGTTTCATAGGATATGATCTTCCACTATAATCATCTTCATCTGC 678
Qy 301 TyrAlaAlaIleIle----- 305
Db 679 TATTCTCAGATCTGCTGCAAACTCTTCAGAACTGCCAAACAAACCCCACTCACTGAGAAA 738
Qy 306 -----ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
Db 739 TCTGGTGTAAACAAAAGGCTCTCAACAAAT----- 771
Qy 321 AlaSerLeuLeuIleValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
Db 772 -----ATTCTTATTATTGTTGTTGTTCTCTGTTTCCACACCTTACCATTGTTGCAAT 825
Qy 341 IleIleHis-----HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 826 ATTCAACATATGATTAAAGAGCTTCGTTCTCTAATTTCTTGGAAATGAGCCAAACACAT 885
Qy 355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db 886 TCGTTCCAGATTTCTCTGCACCTTTACAGTATGCTGATGAACCTTCAATTGCTCATGGAC 945
Qy 374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
Db 946 CTTTATTACTTCTTTCATGTATAA-----GGGTATAAGAGAAAG 987
Qy 394 *****AsnAspLeuArgGluGln 401
Db 988 GTTATGAGGATGCTGAACGGCAA 1011
```

RESULT 20

```
US-10-225-567A-303
; Sequence 303, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
```

;; PRIOR FILING DATE: 2000-12-19
;; NUMBER OF SEQ ID NOS: 2292
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 303
;; LENGTH: 2070
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-225-567A-303

Alignment Scores:
Pred. No.: 1,99e-28 Length: 2070
Score: 410.50 Matches: 107
Percent Similarity: 44.54% Conservative: 48
Best Local Similarity: 30.75% Mismatches: 132
Query Match: 19.03% Indels: 61
DB: 12 Gaps: 12

US-09-208-629F-6 (1-408) x US-10-225-567A-303 (1-2070)

```
QY 62 AlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 177 GCTCCCCAGGCTGATCACCACCTCTCCCTGGCCACGGCAGAG-----221
QY 82 ThrIleThrValLysIleCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 222 -----CAATGTGCCAGGAG-----236
QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeu 121
Db 237 -----ACGCCACTGGGAACATCTGTTCGCCCTCTTACCTT 275
QY 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheAsp 141
Db 276 CTGGATTATCTCGCTTAGTTGGCAATACCTGGCTCTGTGG---CTTTTCATCCGA 332
QY 142 ThrArgSerIleCysThr-----ThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db 333 GACCACAAAGTCGGGACCCGGCCAAACGTTCTCTGATGCATCTGGCCGCGGACTTG 392
QY 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpVal 179
Db 393 TCGTGGCTGCTGCTGCTGCCACCCGCTGGTCTACCACTTCTCTGGGAACCACTGGCCA 452
QY 180 PheGlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db 453 TTTGGGAAATCCATGCCGTCTCACGGCTTCTCTTACCTCAACATGTACGCCACG 512
QY 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 513 ATCTACTTCTACCTGCATCAGCGCCGACCGCTTCTTGCCATTGTGCACCGGTCAG 572
QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 573 TCCTCAAGCTCGCGACGCCCTCTACGCACACATCTGCTGCTTCTGTGGTGGTG 632
QY 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysIleGluTyrTyrLeuValGlnPro 259
Db 633 GTGGCTGTGGCCATGGCCCGCTGTGGTGAGCCCAAG-----671
QY 260 AspileThrThrCysHisAspValHisAsnThrCysGluSerSerSerPropheGlnLeu 279
Db 672 ACCGTGCAGACCAACACACAGGTC-----GTCTGC-----CTGCAGCTG 710
QY 280 Tyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db 711 TACCGGGAGAGCGCTCCACCATGCTGGTGTCTCTCCCTGGCA---GTGGCTTTCACCTTC 767
QY 293 PropheValLeuIleIleTyrCysTyrAlaAlaIleIleIleArgThr-----Leu 308
Db 768 CCGTTATACACAGGTCACCTGCTACCTGCTGATATCCGAGCTGCGGAGGCGCTG 827
QY 309 AsnAlaTyrAspHisArgTrpLeuTrpTyrTyrValLysAlaSerLeuLeuValIle 328
Db -----
```

```
Db 828 COTGTGGAGAACGCGCTCAAGACCAAGCAGTGCATGTCATAGTGTGCGCATC 887
QY 329 PheThrIleCysPheAlaProSerAsnIle-----IleLeuIleIleHisAla 345
Db 888 TTCCTGGTCTGCTTCTGCTGCTTACCACTCAACCGCTCGGTCTAGTGTGCTACCGC 947
QY 346 AsnTyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 948 AGCCATGGGCTCTCTGCGCCACCCAGCGCATCTCTGGCCCTGGCAACCGCATCACCTCC 1007
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPheLeuMetSerLysThr 383
Db 1008 TGCCTCACAGCCTCAACGGGCACTCGACCCCATCATGTATTTCTGCTGCTGAGAAG 1067
QY 384 ArgAsnHisSerThrAlaTyrLeu 391
Db 1068 TTCGCCACGCGCTGTGCAACTTG 1091
```

RESULT 21

US-10-225-567A-218
; Sequence 218, Application US/10225567A
; Publication No. US20030113798A1

GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR;
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 218
; LENGTH: 2424
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-218

Alignment Scores:

Pred. No.:	1,7e-27	Length:	2424
Score:	401.50	Matches:	123
Percent Similarity:	46.62%	Conservative:	70
Best Local Similarity:	29.71%	Mismatches:	161
Query Match:	18.61%	Indels:	60
DB:	9	Gaps:	15

US-09-208-629F-6 (1-408) x US-10-225-567A-218 (1-2424)

```
QY 20 MetLysAlaLeuIlePheAlaAlaGlyLeu-LeuLeuLeuLeuProThrPheCysGln 39
Db 56 ATGTCTGTGCGCCCTGCGCGCGCTGCTGCGCGCTCTCTACCCCTCGGAGCCGC 115
QY 39 nSerGlyMetGluAsnAsp---ThrAsnAsnLeuAlaLysProThrLeuProIleLysTh 58
Db 116 CGCTTAAGTCGAGGAGGAGAGATGACCGAGGTGTGTGGCCGCTGTCTCCCAACGGGAC 175
QY 58 r-----PheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLe 75
Db 176 GCACGCTGCTCTCTCTGCGCGGTCCGGTTCGTC-----209
QY 75 uGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSerAlase 95
Db 210 -----TGGGGAAACAGCAGCGTC-----GCCTC 232
QY 95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr---SerSerLeuSerThrLy 114
Db 233 CACTGCCGCGCTCTCTCTGCTGCTCAATGCGCTTGACCAAGACGGGCTTCCAGTTT 292
QY 114 sLeuIleProAlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaValTh 134
Db -----
```

Db 293 CTACCTGCCGGTGTCTACATCTTGGTATTTCATCGGCTTCTCGGCAACAGCGTGGC 352
Qy 134 rLeuTrpMetLeuPhePheArgThrArgSerIleCys---ThrThrValPheThrAs 153
Db 353 CATCTGGATGTTCTCTCCACATGAAGCCCTGGAGCGCATCTCCGTTGACATGTTCAA 412
Qy 153 nLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheIleAlaThrHisLe 173
Db 413 TTTGGCTCTGGCCGACTCTTGTACGTTGCTGACTCTGCCAGCCCTGATCTTCTACTT 472
Qy 173 uAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTy 193
Db 473 CAATAAACAAGACATGATCTTCGGGATGCGATGTAATCTCAGAGGTTCACTTTCA 532
Qy 193 rGlyAsnMetTyrcysSerIleLeuLeuAlaCysIleSerIleAsnArgTyLeuAl 213
Db 533 TGTGAACCTCTATGGCAGCATCTTTTCTGACATGCATCAGTGCACCGGTACAGCG 592
Qy 213 alleValHisProPheThrTyArgGlyLeuProLysHisThrTyAlaLeuValThrCy 233
Db 593 TGTGGTGTACCCCTCAAGTCCCTGGCCGGCTCAAAAAGAAAGATCGCATCTGTATCAG 652
Qy 233 sGlyLeuValTrpAlaThrValPheLeuTyMetLeuProPhePheIleLeuLysGlnI 253
Db 653 CGTGCTGGTGTGCTCATTTGTGTGGTGGGATCTCCCCC-----ATCCTCTTCTACTC 706
Qy 253 uTyTyTyLeuValGlnProAspIleThr---ThrCysHisAspValHisAsnThrCysG 272
Db 707 AGGTACCGGGTCCGCAAAAACAAACCATCACTCTGTAGCACCC-----752
Qy 272 uSerSerSerPheGlnLeuTyTyTyPheIle-----SerLeuAlaPh 287
Db 753 -ACCTCAGACGAGTACTCGCAAGATTATTTTCATCTACAGCATGTGCACGACCGTGCCCAT 811
Qy 287 ePheGlyPheLeuIleProPheValIleIleTyTyCysTyAlaAlaIleIleArgTh 307
Db 812 G-----TTCTGTCTCCCTGGTGTGATTCCTGGGCTGTACGGATTATTTGTGAGAGC 865
Qy 307 rLeuAsnAlaTyAsp-----HisArgTrpLeuTrpTyTyVally 320
Db 866 TTTGATTACAAGATCTGACAACTCTCTCTGAGGAGAAATCGATTACCTGGTA-- 923
Qy 320 sAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLe 340
Db 924 -----ATCATGTACTGACTGTTTGTGTGTCTTACATCCCTTTCATGTATGAA 976
Qy 340 uIleIleHis--HisAlaAsnTyTyTyTyAsnAsn-----ThrAs 353
Db 977 AACGATGAACCTGAGGCGCGGCTTGATTTTCAGACCCCAATGTGTCTTCAATGA 1036
Qy 353 pGlyLeuTyTyPheIleTyLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAs 373
Db 1037 CAGGGTTATGACGATCAGTGTACAGTGTACAGAGTCTAGCAAGTCTCAACAGTTGTGGA 1096
Qy 373 pProPheLeuTyTyPheLeuMetSerLysThrArgAsnHisSerThrAlaTyTyLeuTh 393
Db 1097 CCCCATCTCTATTCTTGGCGGAGATCTTTTCAAGAGGAGACTCTCCCGACCAACAG 1156
Qy 393 s*****AsnAspLeuArgGlnGlyGlnProSerGln 406
Db 1157 GAAAGCTTCTAGAAGAGTGAGGCAAAATTTGCAATCCAAG 1196

RESULT 22
US-10-024-494-7
; Sequence 7, Application US/10024494
; Publication No. US2003004489A1
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; CAO, LIANG
; ROSEN, CRAIG A.
; TITLE OF INVENTION: Human GPR4 G-Protein Coupled Receptor and
; Nucleotides Encoding Same (As Amended)
; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:
; ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/024,494
; FILING DATE: 21-Dec-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,973
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140001/EKS/HCC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 161..1192
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-024-494-7

Alignment Scores:
Pred. No.: 1,028-26 Length: 1301
Score: 389.00 Matches: 105
Percent Similarity: 47.6% Conservative: 60
Best Local Similarity: 30.3% Mismatches: 130
Query Match: 18.0% Indels: 51
DB: 9 Gaps: 11

US-09-208-629F-6 (1-408) x US-10-024-494-7 (1-1301)
Qy 77 GlyTyTrpThrGlyAlaThrIleThrVallyIleLysCys-ProGluGluSerAlaSerHi 96
Db 76 GGTATTAGAGTTTATAATCTGAATCCCAAGGAGACTCGACGTGATGAAGTGTCTCCAA 135
Qy 96 sLeuHisVal-----LysAsnAlaThrMetGlyTyTrpLeuTh 108
Db 136 ACTGAATTTGACGCTGCTTTACGATGTAACGCTTAACAGCTCCCACTGCTTATAA 195
Qy 108 rSerSerLeuSerThrLysLeuIleProAlaIleTyLeuLeuValPheValGlyVa 128
Db 196 TGACTCCTTTAAGTACACTTTGTATGGGTGTCATGTTTCAGCATGTTGTTGTCTGGTT 255
Qy 128 lProAlaAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys---- 146
Db 256 AATATCCCAATGTGTGCTATATCATTTTC-----ATTCGCTCCT 297
Qy 147 -----ThrThrValPheTyTrpThrAsnLeuAlaIleAlaAspPheLeuPh 161
Db 298 CAAAGTCCGAATGAAGTACACTACATGATTAATGGAATGTCAGACTGCTTTT 357
Qy 161 eCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGl 181

Db	784	TCGGCGGTCATGGGGCTGCTCTTTGGCGTGCCTGCTGCTCACTCTGTTTGCATATGGA	843
Qy	303	AlaIleIleArgThrLeu-----AsnAlaTyrAspHisArgTrpLeu	316
Db	844	CTCATGGCTCGTGGCTGTATCAGCCCTTGCAGGCTCTGCACAGTGTCTTCTCGGCTC	903
Qy	317	TrpTyrValLysAlaSerLeuLeuVallePheThrIleCysPheAlaProSer	336
Db	904	CGCTCTCTCCGACCATAGCTGTGGTGTGCTACTCTTTGCTGCTCTCTGTCGCTTTC	963
Qy	337	AsnIleIleLeuIleIleHis-----AlaAsnTyrTyrAsn	350
Db	964	CACATCACCGCACCATTTACTCTGCGCAGGCTGTGGAGCTGACTGCCAGTACTG	1023
Qy	351	AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSer	370
Db	1024	AACATT-----GTCAACGTGTCTATAAAGTACTCGGCCCTGGCCAGTCCCAACAGC	1077
Qy	371	CysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyr	390
Db	1078	TGCTTGGATCTCTGTCTCTACTTGTCTACTCTGGG-----GlnGlyGlnProSerGlnArg	407
Qy	391	LeuThrLys*****AsnAspLeuArgGlu-----GlnGlyGlnProSerGlnArg	407
Db	1111	---GACAAATATCGACGTACGCTCGTCTGCTGCTGTGTGTGGTGGCAAGCCCGCCG	1167
Qy	408	Thr 408	
Db	1168	ACG 1170	
RESULT 26			
US-10-225-567A-220			
; Sequence 220, Application US/10225567A			
; Publication No. US20030113798A1			
; GENERAL INFORMATION:			
; APPLICANT: LifeSpan Biosciences			
; APPLICANT: Brown, Joseph P.			
; APPLICANT: Burner, Glenn C.			
; APPLICANT: Roush, Christine L.			
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED			
; FILE REFERENCE: 1920-4-4			
; CURRENT APPLICATION NUMBER: US/10/225,567A			
; CURRENT FILING DATE: 2001-12-19			
; PRIOR APPLICATION NUMBER: 60/257,144			
; PRIOR FILING DATE: 2000-12-19			
; NUMBER OF SEQ ID NOS: 2292			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 220			
; LENGTH: 1118			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-225-567A-220			
Alignment Scores:			
Pred. No.: 2,39e-26 Length: 1118			
Score: 384.00 Matches: 101			
Percent Similarity: 48.04% Conservative: 58			
Best Local Similarity: 30.51% Mismatches: 122			
Query Match: 17.80% Indels: 50			
DB: 9 Gaps: 11			
US-09-208-629F-6 (1-408) x US-10-225-567A-220 (1-1118)			
Qy	91	GlulSerAlaSerHisLeuHisVal-----LysAsnAla	102
Db	3	GATGAAGTCTTCCAAATTCGACGTCCTTTACGATGGTAAAGCTTAAACAGC	62
Qy	103	ThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu	122
Db	63	TCCACTGCTTCTATAATGACTCTTTAAGTACACTTTTATGGTGCATGCTTCAGCATG	122
Qy	123	ValPheValGlyValProAlaAsnAlaValThrTrpMetLeuPhePheArgThr	142


```
Db 123 GTGTTTGCTGGTGTAGTATCAATTGTTGGCCATATACATTTTC----- 170
Qy 143 ArgSerIleCys-----ThrThrValPheTyrThrAsnLeuAla 155
Db 171 -----ATCTGCGTCTCAAAAGTCGAAATGAAATACAACTTACATGATTAAGTGGCA 224
Qy 156 IleAlaAspPheLeuPheCysValThrLeuProPheLeuAlaTyrHisLeuAsnGly 175
Db 225 ATCTGACAGCTGCTTTTGTGTTTACTTTTACCTTCCAGGATT-----TTTACTTCAACA 281
Qy 176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
Db 282 CGGAATTTGGCAATTTGGAGATTTACTTTTGAAGATTTCTGTGATGCTGTTTATACAAAC 341
Qy 196 MetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
Db 342 ATGTACGGAAGCATTTCTTTAAACCTGTATTAGTGTAGATCGATTTCTGGCAATTGTC 401
Qy 216 HisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu 235
Db 402 TACCCATTTTAAGTCAAGACTCTTAAGAACCAAGAAATGCAAGATTTGTTGCACCTGGC 461
Qy 236 ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyr 255
Db 462 GTGTGTTAACTGTGATCGGAGAGTGCACCC-----GCCGTT 500
Qy 256 LeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGluSerSerSer 275
Db 501 TTTGTTTCAGTCTACCCACTCTCAGGTTAACAATGCTCAGAACCTGCTTTGAAATTTT 560
Qy 276 Pro-----PheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287
Db 561 CCAGAACCCACATCGAAACATATCTCTCAAGGATTTGTAATTTTCATCGAAATAGTG--- 617
Qy 288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db 618 ---GGATTTTATTTCTCTAATTTTAAATGTAACTTGTCTAGTAGTGTGCTTAAATACT 674
Qy 308 LeuAsn-----AlaTyrAspHisArgTyrLeuTyrValLysAlaSerLeuLeu 324
Db 675 TTAACCAACACGATTACATTAAGTAGAAGCAAAATAACAAACTAAGGTTTAAANAATG 734
Qy 325 Ile-----LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuLeu 341
Db 735 ATTTTGTATCATTTGATCATATTTCTGTTTCTGTTTCTGTTTCTTACAAATATCAATCTATT 794
Qy 342 IleHis-----HisAlaAsnTyrTyrTyrAsnAsnThr-----AspGlyLeuTyr 356
Db 795 TTATATCTCTGTGAGAACACAAACATTTGTTAATTTGCTTCAGTAGTGGCAGAGTAAG 854
Qy 357 PheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeu 376
Db 855 ACAATGTACCAATCACTCTCTGTTTGTGTTTCCCACTGTTGTTTGGACCTATAGTT 914
Qy 377 TyrPheLeuMetSerLysThrArgAsnHisSer 387
Db 915 TACTACTTTATCATCGGACACAAATTCAGAAATCA 947

RESULT 27
US-10-225-567A-331
; Sequence 331, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
```

```
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 331
; LENGTH: 1098
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-331

Alignment Scores:
Pred. No.: 3,57e-26 Length: 1098
Score: 382.00 Matches: 100
Percent Similarity: 46.82% Conservatives: 47
Best Local Similarity: 31.85% Mismatches: 127
Query Match: 17.71% Indels: 40
DB: 9 Gaps: 10

US-09-208-629F-6 (1-408) x US-10-225-567A-331 (1-1098)

Qy 115 LeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThr 134
Db 109 CTGCTGCCTGTGAGCTATGCACTGTCTTTGTCTGGCTTGGCCCTTAAAGCCCAACC 168
Qy 135 LeuTyrMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThr---Asn 153
Db 169 CTATGCTCTTTCATCTTCGCTCCGACCTGGGATGCAACGGCCACCTACATGTTCCAC 228
Qy 154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
Db 229 CTGGCATTTGTCAGACACCTTTGTATGCTGCTGCTGCCACCCCTCATCTACTATTATGCA 288
Qy 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193
Db 289 GCCCAACACCACTGGCCCTTTGGCCTGAGATCTGCAAGTTCGCTCGCTTTCTTTCTAT 348
Qy 194 GlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
Db 349 TGGAACTCTACTGCACTGCTCTTTCTTCCTCCTCAGCTGAGCGTGCACCGCTACCTGGGC 408
Qy 214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
Db 409 ATCTGCCACCACTTCGGGCATCAGCTGGGGCGCCCTCGCTCGAGGCTTCTCTGC 468
Qy 234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro-----PhePheIleLeuLys 251
Db 469 CTGGCAGATTGGTTGGTGTAGCGCTGCTCGTCCGTCGCAACCTGTTCTTTGTCAACACC 528
Qy 252 GlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCys 271
Db 529 AGCAAC-----AAAGGGACACCGCTCTGTGCCATGACACCACTCGGCCCTGAA 576
Qy 272 GluSerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeu 291
Db 577 GAG-----TTTGACCACTATGTGCACTTCAGCTCGCGGCTCATGGGGCTGCTC 624
Qy 292 -----IleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeu--- 308
Db 625 TTTGGCGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 684
Qy 309 -----AsnAlaTyrAspHisArgTyrLeuTyrTyrValLysAlaSerLeu 323
Db 685 CAGCCCTTTCAGGCTGTCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
Qy 324 LeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHis 343
Db 745 GTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804
Qy 344 His-----AlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPhe 357
Db 805 TACCTGGCCAGGCTGTTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 858
Qy 358 IleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 377
Db 859 GTCTATAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 918
```

QY 378 PheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAsp 397
|||
Db 919 TTGCTACTGGG-----CACAATATCGACGTCAG 948

QY 398 LeuArgGlu-----GlnGlyGlnProSerGlnArgThr 408
|||
Db 949 CTCGGTCAGCTGTGTGTGTGTGCAAGCCGCCCGCCGACG 990

RESULT 28
US-09-828-478-1
; Sequence 1, Application US/09828478
; Patent No. US2002015528A1
; GENERAL INFORMATION:
; APPLICANT: Xisao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: Protein
; CURRENT APPLICATION NUMBER: US/09/828,478
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-828-478-1

Alignment Scores:
Pred. No.: 1,07e-25 Length: 1041
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservativity: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: 9 Gaps: 9

US-09-208-629F-6 (1-408) x US-09-828-478-1 (1-1041)

QY 82 ThrIleThrValLysLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
|||
Db 31 TCATCTCCGTATCAGAAATGGACCAATGGCAGCTTCAGCAATACACACAGGAAAC 90

QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeu 121
|||
Db 91 TGCACAAT-----GAAACTTCAAGAGAGAAATTTTCCCAATGTATATCTG 138

QY 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheArg 141
|||
Db 139 ATAATATTTCTGGGGAGCTTGGGAAATGGGTGTCCATATAT--GTITTCCTGGCAG 195

QY 142 -----ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
|||
Db 196 CCTTATAGAGATCCACATCTGTGAAGTTTTCATGCTAATCTGCCATTCAGATCTC 255

QY 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
|||
Db 256 CTGTTTCAAGACGCGCTCCCTTCAGGGCTGACTATATCTTAGAGGCTCCAATTCGATA 315

QY 180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
|||
Db 316 TTGGAGACTGGCGCTGACGATTATGCTTATCTTGTGTGTCACATGTACACGAGT 375

QY 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
|||
Db 376 ATTATTTCTGACCGGCTGAGTGTGTGCTTCTGCAATGTTTCCACCTTCGG 435

QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
|||
Db 436 CTTCTGCATCCACCATCAGGAGTGGCTGGATCTCTGTGGGATCATATGG----- 489

QY 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
|||
Db 490 ---ATCCCTATCATGGCTTCCTCAATAATGCTCCTGGACAGTGGCTCTCAGCAGAACGGC 546

QY 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
|||
Db 547 AGTGTACATCATGCTTAGAGCTGAAT-----CTCTATAAAATTTGCTAAGCTGACGACC 600

QY 280 TyrTyrPheIleSerIleAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyr 299
|||
Db 601 ATGAACATATATTGCTTGGTG---GTGGGCTCGCTGCTGCCATTTTTCACACTCAGCATC 657

QY 300 CysTyrAlaAlaIleAlaIleArgThrLeuAsnAlaTyrAsp----- 312
|||
Db 658 TGTATCTGCTCATCTTCGGGTCTGTAAAGTGGAGGTCCAGAAATCGGGGCTGCCG 717

QY 313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuValIlePheThr 330
|||
Db 718 GTTCTCACAGGAGCA-----CTGACCACCATCATCATCCTTGATCATCTTCTTC 771

QY 331 IleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsn 350
|||
Db 772 TTGTGTTTCTGCTCATCATCAGAGGACCGTCCACTTGACGACATGGAAGTGGGT 831

QY 351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
|||
Db 832 TTATGCAAGACAGACTGCATAAAGCTTGTGTATACACTGGCTTGGCAGACCAAT 891

QY 370 SerCysLeuAspProPheLeuTyrPhe 378
|||
Db 892 GCCTGCTTCAATCCTCTGCTCTATTATAC 918

RESULT 29

US-09-826-791-5
; Sequence 5, Application US/09826791
; Patent No. US20010039037A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc
; TITLE OF INVENTION: No. US20010039037A1el Polypeptide
; FILE REFERENCE: PC10914ADAM
; CURRENT APPLICATION NUMBER: US/09/826,791
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 0008504.3
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/198,367
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-826-791-5

Alignment Scores:
Pred. No.: 1,07e-25 Length: 1041
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservativity: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: 9 Gaps: 9

US-09-208-629F-6 (1-408) x US-09-826-791-5 (1-1041)

QY 82 ThrIleThrValLysLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
|||
Db 31 TCATCTCCGTATCAGAAATGGACCAATGGCAGCTTCAGCAATACACACAGGAAAC 90

QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeu 121
|||

```
Db 91 TGCACAAATT-----GAAACTTCAAGAGAGAAATTTTCCCAATTTGATATCTG 138
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheAsg 141
Db 139 ATAATAATTTTCGGGGAGTCTTGGGAAATGGGTGTCATATAT---GTTTTCCTGCG 195
Qy 142 -----ThrArgSerIleCysThrValPheTyrThrAsnLeuAlaAlaAspPhe 159
Db 196 CTTTATAAGAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGGCCATTTCCAGATC 255
Qy 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db 256 CTGTTTATAAGCAGCTTCCCTTCAAGGCTGACTATATCTTAGAGGCTCCAATGGATA 315
Qy 180 PheGlyGluValLeuLeuValThrValPheTyrGlyAsnMetTyrCysSer 199
Db 316 TTTGGAGACTGGCTGAGGATATGCTTATCTTATCTTGTATGCTCAATGTCACAGCAT 375
Qy 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaValHisProPheThr 219
Db 376 ATTTATTTCTGACCGTGTGAGTGTGCGCTTTCCTGCAATGTTCCACCCCTTTCG 435
Qy 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 436 CTTCTGCATGTCACAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG----- 489
Qy 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db 490 ---ATCCTTATCTGCTTCTCCTCAATAATGCTGTCGAGTGCCTGAGCAGAGACGGC 546
Qy 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerPheProPheGlnLeu 279
Db 547 AGTGTACATCATGCTTAGAGCTGAAT-----CTCTATAAAATGCTAAGCTGACAGC 600
Qy 280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleValHisProPheThr 299
Db 601 ATGAATATATATGCTTGGT---GTGGGCTGCTGCTGCCATTTTTCACATCAGCATC 657
Qy 300 CysTyrAlaAlaIleAlaIleArgThrLeuAsnAlaTyrAsp----- 312
Db 658 TGTATCTGCTCATCATCTGCGGTTCTGTTAAAAGTGGAGTCCCAAGATCGGGCTCGCG 717
Qy 313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuLeuValIlePheThr 330
Db 718 GTTTCCTCAGGAAGGCA-----CTGACCACCATCATCATCTTGTATCATCTTCTTC 771
Qy 331 IleCysPheAlaProSerAsnIleIleLeuIleHisAlaAsnTyrTyrTyrAsn 350
Db 772 TTGTGTCTTCTGCTTATCACACATGAGGACCGTCCACCTTGACGACATGGAAGTGGGT 831
Qy 351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db 832 TTATGMAAGACAGACTGCATAAAGCTTGTGTTATCATCATCTGCGCTTGGCAGCGCAAT 891
Qy 370 SerCysLeuAspProPheLeuTyrPhe 378
Db 892 GCCTGCTTCAATCTCTGCTCTATATAC 918

RESULT 30
US-09-866-230-6
; Sequence 6, Application US/09866230
; Patent No. US20020150901A1
; GENERAL INFORMATION:
; APPLICANT: Murphy, Andrew, et al.
; TITLE OF INVENTION: No. US20020150901A1el Nucleic Acids, Polypeptides, Methods of Mak
; FILE REFERENCE: REG 771A
; CURRENT APPLICATION NUMBER: US/09/866,230
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/207,725
; PRIOR FILING DATE: 2000-05-26
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 6
; LENGTH: 1041
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1041)
US-09-866-230-6

Alignment Scores:
Pred. No.: 1,07e-25 Length: 1041
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservative: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: 10 Gaps: 9

US-09-208-629f-6 (1-408) x US-09-866-230-6 (1-1041)

Qy 82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 31 TCCATCTCGTATCAGAAATGGAACCAATGCGACCTTCCAGCAATTAACAACAGCAGAAC 90
Qy 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db 91 TGCACAAAT-----GAAAACTTCAAGAGAGAAATTTTCCCAATTTGATATCTG 138
Qy 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheAsg 141
Db 139 ATAATAATTTTCTGGGAGTCTTGGAAATGGGTGTCATATAT---GTTTTCCTGCG 195
Qy 142 -----ThrArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db 196 CTTTATAAGAGTCCACATCTGTGAACGTTTTCATGCTAAATCTGCGCATTTCCAGATCTC 255
Qy 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db 256 CTGTTTATAAGCAGCTTCCCTTCAAGGCTGACTATATCTTAGAGGCTCCAATTTGGATA 315
Qy 180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db 316 TTTGAGAGCTCGCTGCGAGGATATGCTTATCTCTGTATGTCACATGTCACAGCAT 375
Qy 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 376 ATTTATTTCTGACCGTGTGAGTGTGCGCTTTCCTGCAATGTTCCACCCCTTTCG 435
Qy 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 436 CTTCTGCATGTCACAGCATCAGGAGTGCCTGGATCCTCTGTGGGATCATATGG----- 489
Qy 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db 490 ---ATCCTTATCATGCTTCTCAATAATGCTGTCGAGTGCCTGAGCAGAGACGGC 546
Qy 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerPheProPheGlnLeu 279
Db 547 AGTGTACATCATGCTTAGAGCTGAAT-----CTCTATAAAATGCTAAGCTGACAGC 600
Qy 280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleValHisProPheThr 299
Db 601 ATGAATATATATGCTTGGT---GTGGGCTGCTGCTGCCATTTTTCACATCAGCATC 657
Qy 300 CysTyrAlaAlaIleAlaIleArgThrLeuAsnAlaTyrAsp----- 312
Db 658 TGTATCTGCTCATCATCTGCGGTTCTGTTAAAAGTGGAGTCCCAAGATCGGGCTCGCG 717
Qy 313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuLeuValIlePheThr 330
Db 718 GTTTCCTCAGGAAGGCA-----CTGACCACCATCATCATCTTGTATCATCTTCTTC 771
Qy 331 IleCysPheAlaProSerAsnIleIleLeuIleHisAlaAsnTyrTyrTyrAsn 350
Db 772 TTGTGTCTTCTGCTTATCACACATGAGGACCGTCCACCTTGACGACATGGAAGTGGGT 831
Qy 351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db 832 TTATGMAAGACAGACTGCATAAAGCTTGTGTTATCATCATCTGCGCTTGGCAGCGCAAT 891
Qy 370 SerCysLeuAspProPheLeuTyrPhe 378
Db 892 GCCTGCTTCAATCTCTGCTCTATATAC 918
```

Db 772 TTGTTGTTCTCCCTATCACACACTGAGGACCGTCCACTTGACGACATGGAAGTGGGT 831
QY 351 ---AsnThrAspGlyLeuTyrPheLeuLeuAlaLeuCysLeuGlySerLeuAsn 369
Db 832 TTATGCAAGACAGACTGCATAAAGCTTTGGTTATCATCACTGGCCTTGGCAGCAGCAAT 891
QY 370 SerCysLeuAspProPheLeuTyrPhe 378
Db 892 GCCTGCTTCATCCTCTGCTCTATTAC 918

RESULT 31

US-09-779-679-1
; Sequence 1, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; FILE REFERENCE: 15966-661
; CURRENT APPLICATION NUMBER: US/09/779, 679
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/183191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180, 929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (105)..(1142)
; NAME/KEY: misc feature
; LOCATION: (120)..(122)
; OTHER INFORMATION: Alternative ATG start site
US-09-779-679-1

Alignment Scores:
Pred. No.: 1.4e-25 Length: 1260
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservative: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: 9 Gaps: 9

US-09-208-629F-6 (1-408) x US-09-779-679-1 (1-1260)
QY 82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 135 TCATCTCCGTATCAGAAATGCAACATGGCACCTTCAGCAATAAACAACAGAGAAC 194
QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLeuLeuProAlaIleTyrLeu 121
Db 195 TGCACAAAT-----GAAACTTCAAGAGAGAAATTTTCCCAATGTTATATCTG 242
QY 122 LeuValPheValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPheArg 141
Db 243 ATAATATTTTCTGGGAGTCTTGGAAATGGTTGTCCATATAT---GTTTCTCTGCAG 299
QY 142 -----ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaaspPhe 159
Db 300 CCTTATAAGAAGTCCACATCTGTGAACGTTTTCATGTAAATCTGGCCATTTTCAGATCTC 359
QY 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db 360 CTGTTCAATAGCACGCTTCCCTCAGGCTGACTATATCTTAGAGGCTCCCAATTGGATA 419
QY 180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
Db 420 TTTGAGACCTGGCCTGCAGGATTATGCTTATTCCTTGTATGTCAACATGTACAGCAGT 479
QY 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 480 ATTTATTTCTCCGACCGTCTGAGTGTGCGTTTCTGGCAATGGTTCACCCCTTCGG 539
QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 540 CTTCTGCATGCACAGCATCAGGAGTGGTGGATCTCTCTGGGATCATATGG-----593
QY 240 ValPheLeuTyrMetLeuProPheIleLeuLysGlnIuTyrTyrLeuValGlnPro 259
Db 594 ---ATCCTTATCATGGCTTCTCAATAATGCTCTGGACAGTGGCTCTGAGCAGAAGCGC 650
QY 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db 651 AGTGTCAATCATGCTTAGAGCTGAAT-----CTCTATAAAATTTGCTAAGCTGCAGACC 704
QY 280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyr 299
Db 705 ATGAACATATATGCTTGGTG---GTGGGCTGCTGTGCCATTTTTCACACTCAGCATC 761
QY 300 CysTyrAlaAlaIleAlaIleArgThrLeuAsnAlaTyrAsp-----312
Db 762 TGTATCTGCTCATCATTCGGGTTCTGTTAAAAGTGGAGTCCCAAGATCGGGGTGCGG 821
QY 313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuValIlePheThr 330
Db 822 GTTCTCACAGGAAGCA-----CTGACCACCATCATCATCCTTGATCATCTTCTTC 875
QY 331 IleCysPheAlaProSerAsnIleIleLeuIleHisHisAlaAsnTyrTyrTyrAsn 350
Db 876 TTGTGTTTCTGCTCCCTATCACACTGAGGACCGTCCACTTCAGCAGATGGAAGTGGGT 935
QY 351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db 936 TTATGCAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCCTTGGCAGCAGCAAT 995
QY 370 SerCysLeuAspProPheLeuTyrPhe 378
Db 996 GCCTGCTTCATCCTCTGCTCTATTAC 1022

RESULT 32

US-09-779-679-24
; Sequence 24, Application US/09779679
; Publication No. US20030082757A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond J
; APPLICANT: Burgess, Catherine E

```
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Mishra, Vishnu S
; APPLICANT: Casman, Stacie
; APPLICANT: Ballinger, Robert
; APPLICANT: Vernet, Corine A
; APPLICANT: Li, Li
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Andrew, David P
; APPLICANT: Mezes, Peter S
; TITLE OF INVENTION: No. US20030082757A1el Proteins and Nucleic Acids Encoding the Sam
; CURRENT APPLICATION NUMBER: US/09/779, 679
; CURRENT FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181045
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/181191
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: USSN 60/180, 929
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: USSN 60/219758
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181339
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/181344
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/221341
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: USSN 60/181392
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: USSN 60/219585
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 24
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-779-679-24

Alignment Scores:
Pred. No.: 1,4e-25 Length: 1260
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservative: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: Gaps: 9

US-09-208-629f-6 (1-408) x US-09-779-679-24 (1-1260)

QY 82 ThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsn 101
Db 135 TCATCTCCGTATCAGAAATGGAACCAATGGCACCTTCAGCAATAACACAGCAGGAAC 194
QY 102 AlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeu 121
Db 195 TGCACAAAT-----GAAACTTCAGAGAGAAATTTTCCCAATGTATATCTG 242
QY 122 LeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg 141
Db 243 ATAATATTTTCTGGGGAGTCTTGGGAATGGGTTCCTCATATAT--GTTTCTCTGCAG 299
QY 142 -----ThrArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
Db 300 CCTATAGAAGGCCATCTGTGAACGTTTTCATGCTAAATCTGCGCAATTCAGATCTC 359
QY 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVal 179
Db 360 CTGTTCAATAGCAGCTTCCCTCAGGGCTGACTATATCTTAGAGCTCCAAATGGATA 419
QY 180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
```

```

420 TTTGGAGACCTGGCCTGCAGGATTATGCTTATTCTTATCTGTATCTCAACATGTACAGCACT 479
QY 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 480 ATTTATTTCTCAGCGTGTGAGTGTGTGGCTTTCTTGGCAATGGTTTCCACCCCTTCGG 539
QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr 239
Db 540 CTCTCTGCATGTCACCAAGCATCAGGAGTGCCTGATCTCTGTGGGATCATATGG----- 593
QY 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db 594 ---ATCCTTATCATGCTTCTCAATAATGCTCTGGACAGTGGCTCTGAGCAGAACGGC 650
QY 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu 279
Db 651 AGTGTCAATCATGCTTAGAGCTGAAT-----CTCTATAAAATGCTAAGCTGCAGACC 704
QY 280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyr 299
Db 705 ATGAACTATATTGCTTGGTG--GTGGGCTGCTGCTGCCATTTTTCACACTCAGCATC 761
QY 300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAsp----- 312
Db 762 TGTATCTGCTGATCATTCGGGTTCCTTTAAAGTGGAGGTCCACAGAAATCGGGCTGCG 821
QY 313 -----HisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThr 330
Db 822 GTTCTCTCAGAGGAGCA-----CTGACCACCATCATCATCCTTGATCATCTTCTTC 875
QY 331 IleCysPheAlaProSerAsnIleIleLeuIleHisHisAlaAsnTyrTyrAsn 350
Db 876 TTGTGTTTCTGCTCCCTATCACACATCAGGACCGTCCACTTCACGACATGGAAGTGG 935
QY 351 ---AsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsn 369
Db 936 TTATGCAAGACAGACTGCATAAAGCTTTGGTTATCACACTGGCTTGGCAGCAGCAAT 995
QY 370 SerCysLeuAspProPheLeuTyrPhe 378
Db 996 GCCTGCTTCAATCCTCTGCTCTATTAC 1022

RESULT 33
US-09-828-478-3
; Sequence 3, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; TITLE OF INVENTION: Protein
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828, 478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 3
; LENGTH: 1430
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-828-478-3

Alignment Scores:
Pred. No.: 1.68e-25 Length: 1430
Score: 376.50 Matches: 93
Percent Similarity: 50.16% Conservative: 62
Best Local Similarity: 30.10% Mismatches: 129
Query Match: 17.45% Indels: 25
DB: Gaps: 9
```


US-09-826-791-1

Alignment Scores:

Pred. No.: 1.54e-25 Length: 993
Score: 374.50 Matches: 91
Percent Similarity: 50.17% Conservative: 60
Best Local Similarity: 30.23% Mismatches: 125
Query Match: 17.36% Indels: 25
DB: 10 Gaps: 9

US-09-208-629F-6 (1-408) x US-09-826-791-1 (1-993)

QY 90 ProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSer 109
DB 7 CCAATGGCACCTTCAGCAATAACAACAGCAGGAAGTGCACATT-----GAA 54
QY 110 SerLeuSerThrLysLeuLeuProAlaIleTyrLeuLeuValPheValValGlyValPro 129
DB 55 AACTTCAAGAGAGAAATTTTCCCAATGTATATCTGATAATATTTTCTGGGAGCTGTG 114
QY 130 AlaAsnAlaValThrLeuTrpMetLeuPhePheArg-----ThrArgSerIleCysThr 147
DB 115 GGAATGGGTGTGCCATATAT---GTTTCTGCAGCCTTATAAGAAAGTCCACATCTGG 171
QY 148 ThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPhe 167
DB 172 AAGTTTTCATGCTAATCTGGCCATTTCAGATCTCTGTTCATAAGCAGCGTCCCTTC 231
QY 168 LysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAla 187
DB 232 AGGCTGACATATATCTTAGAGGCTCCAATGGATATTTGGAGACCTGGCCTGCAGATT 291
QY 188 ThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSer 207
DB 292 ATGCTTATTCCTGTATGTCAACATGTACAGCAGTATTTATTTCTCCCGCTGCAGT 351
QY 208 IleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThr 227
DB 352 GTTGTGGTTCCTGGCAATGGTTACCCCTTTTCGGCTTCTGCATGTCACAGCATCAGG 411
QY 228 TyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhe 247
DB 412 AGTGCCTGATCTCTGTGGATCATATG-----ATCCTTATCATGCTTCTCTCA 462
QY 248 PheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspVal 267
DB 463 ATAATGCTCTGGACAGTGGCTCTGAGCAGAACGGCAGTGTCAATCATGCTTAGAGCTG 522
QY 268 HisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhe 287
DB 523 AAT-----CTCTATAAATTTGTAAGCTGACGACCATGAACTATATTTGCTTGGTG 573
QY 288 PheGlyPheLeuLeuProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
DB 574 GTGGGCTGCTGCTGCCAATTTTTCACACTCAGCATCTGTATCTGTGCTGATCATTCGGTT 633
QY 308 LeuAsnAlaTyrAsp-----HisArgTrpLeuTrpTyr 318
DB 634 CTGTTAAAGTGAGGTGCCAGAAATCGGGGCTCGGGGTTCTCACAGGAAGCA----- 687
QY 319 ValLysAlaSerLeuLeuValIlePheThrIleCysPheAlaProSerAsnIle 338
DB 688 CTGACCACCATCATCATCATCTTGTATCACTTCTTCTGTGTTTCTGCTCCCTATCACACA 747
QY 339 IleLeuIleIleHisIleAsnTyrTyrTyrAsn---AsnThrAspGlyLeuTyrPhe 357
DB 748 CTGAGGACCGCTCCACTTCAGCATGGAAAGTGGGTTTATGCAAGACAGACTGCATAAA 807
QY 358 IleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyr 377
DB 808 GCTTTGGTTATACACTGTGGCCTTGGCAGCAGCAATGCTGCTTCAATCCTCTGCTCTAT 867
QY 378 Phe 378

DB 868 TAC 870

RESULT 37

US-10-225-567A-546
; Sequence 546, Application US/10225567A
; Publication No. US20030113798A1

GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIORITY FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIORITY FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 546
; LENGTH: 1014
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-546

Alignment Scores:

Pred. No.: 2.42e-25 Length: 1014
Score: 372.50 Matches: 103
Percent Similarity: 47.51% Conservative: 59
Best Local Similarity: 30.21% Mismatches: 140
Query Match: 17.27% Indels: 39
DB: 9 Gaps: 12

US-09-208-629F-6 (1-408) x US-10-225-567A-546 (1-1014)

QY 92 GluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeu 111
DB 4 CATGAAACAGGAATCTGACAGTATCTTCTGCCACATGCCATGACACATATTGATGATC 63
QY 112 SerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAlaAsn 131
DB 64 CGCAATCAAGTGTATTCACCTTGTACTCTATGATCTCTGTGTAGGCTTCTTTGGCAAT 123
QY 132 AlaValThrLeuTrpMetLeuPhe-----PheArgThrArgSerIleCysThrThrVal 149
DB 124 GCGTTTGTGCTCTATGTCTCTATAAAACCTTATCAAGAAAGTCA---GCCTTCCAAGTA 180
QY 150 PheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIle 169
DB 181 TACATGATTAATTTAGCAGTAGCAGATCTACTTTGTGTGTGCACACTGCCTCTCCGTGTG 240
QY 170 AlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThr 189
DB 241 GTCTATTATGTTCAAAAGGCAATTTGGCTCTTTGGTGACTTCTTGTGGCGCTCAGCACC 300
QY 190 ValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsn 209
DB 301 TATGCTTTGTATGTCAACCTCTATTGTAGCATCTTCTTTATGACAGCCATGAGCTTTTTC 360
QY 210 ArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAla 229
DB 361 CGGTGCAITGCAATTTGTTTCCAGTCCAGAACATTAATTTGGTTACACAGAAAAAGCC 420
QY 230 LeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIle 249
DB 421 AGGTTTGTGTGTAGGATTTTGGATTTTGTGATTTTGTGACCATTTCTCAATTTCTAATG 480
QY 250 LeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsn 269
DB 481 GCCAAACACAAA-----AAAGATGAGAAAAATAATACCAAGTGTCTTTGAGCCCCCACA 534
QY 270 ThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGly 289


```
Db 535 GACAATCAAACTAAATCAATGTTGGTCTTGCAATTATGTGTCATG---TTTGTGGC 591
Qy 290 PheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleAArgThrLeu--- 308
Db 592 TTTATCATCCCTTTTGTATTATAATTTGCTGTTACAAATGATCATTTTGACCTTACTA 651
Qy 309 -----AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSer 322
Db 652 AAAAAATCAATGAAAAAATCTGTCAAGTCATAAA-----AAGGCTATA 696
Qy 323 LeuLeuIleLeuValIle-----PheThrIleCysPheAlaProSerAsnIleIle 339
Db 697 GGAATGATCAAGTCGCGACCGTCCCTTTTATGTCAGTTTCATGTCATATCATATTCAA 756
Qy 340 LeuIleIleHisAlaAsnTyrTyrAsnAsnThr-----AspGlyLeuTyr 356
Db 757 CGTACCATT---CACCTTCATTTTACACAATGAACTAAACCTGTGATTCGTCCTT 813
Qy 357 PheIle-----TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db 814 AGAATGAGAAGTCCGTTGGTGCATAAACCTTGTCTGTGCTGCATCCCAATTGCTTTGAC 873
Qy 374 ProPheLeuTyrPheLeu-----MetSerLysThrArgAsn 395
Db 874 CCTCTCTATATTCTTTTCTGGGGTAACTTTAGGAAAGCGCTGTCTACATTCAGAAAG 933
Qy 386 His-----SerThrAlaTyrLeuThrLys*****AsnAspLeuArgGluGlnGly 402
Db 934 CATCTTTGTCAGCGTGACTTAATGACCAGAAAGAGCCCTCTTTGCCAGAAAGGA 993
Qy 403 Gln 403
Db 994 GAA 996
```

RESULT 38

```
US-10-167-192-4
; Sequence 4, Application US/10167192
; Publication No. US20030040052A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: No. US20030040052A1el G-Protein Coupled Receptors
; FILE REFERENCE: 5800-11a
; CURRENT APPLICATION NUMBER: US/10/167,192
; CURRENT FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: US/09/420,187
; PRIOR FILING DATE: 1999-10-18
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1358
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-167-192-4
```

```
Alignment Scores:
Pred. No.: 5,04e-25 Length: 1358
Score: 371.00 Matches: 106
Percent Similarity: 46.41% Conservative: 62
Best Local Similarity: 29.28% Mismatches: 144
Query Match: 17.20% Indels: 50
DB: 9 Gaps: 13
```

```
US-09-208-629F-6 (1-408) x US-10-167-192-4 (1-1358)
```

```
Qy 82 ThrIleThrValLysIleLysCysProGluSerAlaSer----- 95
Db 170 ACATTTGTGAAGTACTCCAGTCCAGAAAGAGCGCAAGACGACATTCGTAGAGAAAC 229
Qy 96 -----HisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSer 110
Db 230 ATGGATGAACAGGAATCTGACAGTATCTTCTGCCACATGCCCATCTATTGTATGAC 289
```

RESULT 39

```
US-09-826-508-9
; Sequence 9, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
```

```
Qy 111 LeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGlyValProAla 130
Db 290 TTCCGCAATCAAGTGTATTCCACCTTGTAATCTATGATCTCTGTTGTAGGCTTCTTTGGC 349
Qy 131 AsnAlaValThrLeuTrpMetLeuPhe-----PheArgThrArgSerIleCysThrThr 148
Db 350 AATGGCTTTGTGCTCTATGCTCTATATAAAACCTATACAAAGAGTCA---GCCTTCCAA 406
Qy 149 ValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLys 168
Db 407 GTATACATGATTAAATTAGCAGTAGCAGATCTACTTTGTGTGTGACACACTGCTCTCCGT 466
Qy 169 IleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThr 188
Db 467 GTGGCTATTATATTTACAAAGGCATTTGGCTCTTTGTGTGACTTCTTTGTGCGCCTCAGC 526
Qy 189 ThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIle 208
Db 527 ACCTATGCTTTGTATGTCACCTCTATTGTAGCATCTCTTTATGACACCCATGAGCTTT 586
Qy 209 AsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyr 228
Db 587 TTCCGGTGTGATGCAATTGTTTTCAGTCCAGACACATAATTGTTTACACAGAAAAA 646
Qy 229 AlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePhe 248
Db 647 GCCAGTTTGTGTGTAGTATTTGGATTTTGTGATTTTGCACAGTTCTCCATTTCTA 706
Qy 249 IleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHis 268
Db 707 ATGGCCAAACCAACAA-----AAAGATGAGAAAAATAATACCAAGTCTTTGAGCCCCA 760
Qy 269 AsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePhe 288
Db 761 CAAGACAATCAAACTAAAAATCATGTTTGGTCTTGCAATATGTGTCATGTTGTT---GTT 817
Qy 289 GlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleAArgThrLeu 308
Db 818 GGCTTTATCATCCCTTTTGTATTATATTTGTTGTTACACAATGATCATTTTGACCTTA 877
Qy 309 -----AsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAla 321
Db 878 CTAAAAAATCAATGAAAAAATCTGTCAAGTCATAAA-----AAGGCT 922
Qy 322 SerLeuLeuIleLeuValIle-----PheThrIleCysPheAlaProSerAsnIle 338
Db 923 ATAGGAATGATCATGTCGTGACCCGCTGCTTTTGTAGTCAGTTTATGCAATATCATATT 982
Qy 339 IleLeuIleIleHisAlaAsnTyrTyrAsnAsnThr-----AspGlyLeu 355
Db 983 CAACGTACCATT---CACCTTCATTTTACACAATGAAACTAAACCCCTGTGATTCGTC 1039
Qy 356 TyrPheIle-----TyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeu 372
Db 1040 CTTAGAATGACAGAGTCGCTGTCATAACCTTGTCTGTGCTGCATCCCAATTTGTTGCTTT 1099
Qy 373 AspProPheLeuTyrPheLeu-----MetSerLysThrArg 384
Db 1100 GACCTCTCTATATTTCTTTTCTGGGGTAACTTTTAGGAAAGCGTGTCTACATTTAGA 1159
Qy 385 AsnHis-----SerThrAlaTyrLeuThrLys*****AsnAspLeuArgGluGln 401
Db 1160 AAGCATTTCTTGTCCAGCGTGACTTATGTACCAGAAAGAGCGCTCTTTGCCAGAAAAA 1219
Qy 402 GlyGln 403
Db 1220 GGAGAA 1225
```

APPLICANT: Nabil Elshourbagy
APPLICANT: Lisa Vawter
TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
TITLE OF INVENTION: and Polynucleotides

FILE REFERENCE: GP-70744USB
CURRENT APPLICATION NUMBER: US/09/826,508
CURRENT FILING DATE: 2001-04-05
NUMBER OF SEQ ID NOS: 40
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9

LENGTH: 1077
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-826-508-9

Alignment Scores:
Pred. No.: 6,2e-25 Length: 1077
Score: 368.50 Matches: 89
Percent Similarity: 50.33% Conservative: 64
Best Local Similarity: 29.28% Mismatches: 116
Query Match: 17.08% Indels: 35
DB: 10 Gaps: 9

US-09-208-629F-6 (1-408) x US-09-826-508-9 (1-1077)

QY 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIle 119
DB 94 AAGAACACCACC-----CTTCACATGAATTTGACACAAATGCTTCGCGGTGCTT 144
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
DB 145 TATCTCATTATTTTGGGCAAGCATCTTCTGAATGGTTAGCAGTGIGG---ATCTTC 201
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
DB 202 TTCACATTAGGAATAAACACAGCTTCATATCTCAAAACACATAGTGGTTGCAGAC 261
QY 159 PheLeuPheCysValThrLeuProPheLysAlaTyrHisLeuAsnGlyAsnAsnTrp 178
DB 262 CTCAATGACGCTGACATTTCCATTTGAAATAGTCATGATCGAGATTGGACCTTGG 321
QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
DB 322 TACTTCAAGTTTATCTCTCAGATACACATTCAGTTTGTGTTTATGCAACATGTATACT 381
QY 199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProphe 218
DB 382 TCCATCGTGTCTCTGGGCTGATAAGCAATTCATCTGCTATCTGAAGGTGGTCAAGCCATT 441
QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTtpAla 238
DB 442 GGGGACTCTCGGATGTACAGCATAACTTTCACGAAGGTTTATCTGTTGTTGGGTG 501
QY 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
DB 502 ATCATGGCTGTTTGTCTTGGCAACATCATCTGACAAATGGT-----CAG 549
QY 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerPropheGln 278
DB 550 CCA-----ACAGAGGACAAATATCCATGACTGTCTCAAAAGCTTAAAGTCTCTGGGG 600
QY 279 LeuTyrTyrPheIleSerLeuAlaPhePheGly-----PheLeuIleProPheVal 295
DB 601 GTCAATGGCATGCGGAGTACCTATGTGTAACAGCTGTGTTTGTGGCCGTGCTGGTG 660
QY 296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
DB 661 ATTCTGATCGATGTTACATAGCCATATCCAGGTACATCCACAAATCCAGGCAATTC 720
QY 316 Leu-----TTPTrpValLysAlaSerLeuLeuLeuValIle 328
DB 721 ATAAGTCAGTCAAGCCGAAACGAAACATAACACAGCATCATCGGTTGTTGTGGCTGTG 780

QY 329 PheThrIleCysPheAlaProSerAsnIleLeuLeuIleHisAlaAsnTyrTyr 348
DB 781 TTTTTCCTGCTTTCTACCATATCATTGTCAGAAAT-----CCTTTACT 828
QY 349 TyrAsnAsnThrAspGly-----LeuTyrPheIleTyr 359
DB 829 TTTAGTCACTTAGACAGGCTTTTAGATGAATCGCACAAAATCCTATATTACTGCAAA 888
QY 360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPropheLeuTyrPheLeu 379
DB 889 GAAATTACACTTTTCTGTCTGCGTGAATGTTGCTGGATCCCAATAATTTACTTTTC 948
QY 380 MetSerLysThr 383
DB 949 ATGTGTAGGTCA 960

RESULT 40

US-10-225-567A-417
; Sequence 417, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 417
; LENGTH: 1402
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-417

Alignment Scores:
Pred. No.: 8,98e-25 Length: 1402
Score: 368.50 Matches: 89
Percent Similarity: 50.33% Conservative: 64
Best Local Similarity: 29.28% Mismatches: 116
Query Match: 17.08% Indels: 35
DB: 9 Gaps: 9

US-09-208-629F-6 (1-408) x US-10-225-567A-417 (1-1402)

QY 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIle 119
DB 343 AAGAACACCACC-----CTTCACATGAATTTGACACAAATGCTTCGCGGTGCTT 393
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
DB 394 TATCTCATTATTTTGGGCAAGCATCTTCTGAATGGTTAGCAGTGIGG---ATCTTC 450
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
DB 451 TTTCCACATTAGGAATAAACACAGCTTCATATCTCAAAACACATAGTGGTTGCAGAC 510
QY 159 PheLeuPheCysValThrLeuProPheLysAlaTyrHisLeuAsnGlyAsnAsnTrp 178
DB 511 CTCCAATGACGCTGACATTTCCATTTGAAATAGTCATGATCGAGATTGGACCTTGG 570
QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
DB 571 TACTTCAAGTTTATCTCTGAGATACACTTCAGTTTGTGTTTATGCAACATGTATACT 630
QY 199 SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProphe 218
DB 631 TCCATCGTGTCTCTGGGCTGATAAGCAATTCATCTGCTATCTGAAGGTGGTCAAGCCATT 690

```
QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 691 GGGGACTCTCGGATGTACAGCATAACCTTCACGAAGGTTTATCTGTTGTTGGGTG 750
QY 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 751 ATCATGGCTGTTTGTCTTGGCAACATCATCTGACAAATGGT-----CAG 798
QY 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db 799 CCA-----ACAGAGGACAATATCCATGACTGCTCAAAACTTAAAGTCCTTTGGGG 849
QY 279 LeuTyrTyrPheIleSerLeuAlaPhePheGly-----PheLeuIleProPheVal 295
Db 850 GTCAATGGCATACGGCAGTCACCTATGTGAACAGCTGCTTGTGTCGGCGTGTGGTG 909
QY 296 LeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrp 315
Db 910 ATTCTGATCGATGTTACATAGCCATATCCAGGTACATCCACAAATCCACGAGGCAATTC 969
QY 316 Leu-----TrpTyrValLysAlaSerLeuLeuIleLeuValIle 328
Db 970 ATAAGTCAGTCAAGCCGAAAGCGAAACATAACACAGAGCATCAGGGTGTGTTGGGTG 1029
QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisIleAlaAsnTyrTyr 348
Db 1030 TTTTTCCTGCTTTCTACCATATCATCTGTGCAGAAAT-----CCTTTTACT 1077
QY 349 TyrAsnAsnThrAspGly-----LeuTyrPheIleTyr 359
Db 1078 TTTAGTCACTTAGACAGGCTTTTAGATGAATCTGCACAAAAATCCTATATTACTGCAA 1137
QY 360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
Db 1138 GAAATTACACTTTTCTTGTCTGCGTGTAAATGTTGCTGGATCCAAATAATTACTTTTC 1197
QY 380 MetSerLysThr 383
Db 1198 ATGTGTAGGTCA 1209
```

Search completed: June 29, 2003, 11:46:58
Job time : 174.162 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model.

Run on: June 29, 2003, 07:52:59 ; Search time 46.0564 Seconds
(without alignments)
2716.756 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILOISXRLRDGTQVIMK.....AYLTXXNDLRQGQPSORT 408

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spoal/US09208629/runat_24062003_101626_12944/app_query.fasta_1.1166
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CEN 1.1 47 @runat_24062003_101626_12944 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	2142	99.3	1224 2 US-08-742-440A-4 Sequence 4, Appli
2	1947	90.3	1102 2 US-08-742-440A-5 Sequence 5, Appli
3	1443	66.9	1224 2 US-08-742-440A-1 Sequence 1, Appli
4	959	44.5	1224 2 US-08-742-440A-2 Sequence 2, Appli
5	606.5	28.1	2732 1 US-08-476-000-60 Sequence 60, Appli
6	606.5	28.1	2732 1 US-08-476-000-60 Sequence 60, Appli
7	606.5	28.1	2732 1 US-08-476-000-60 Sequence 60, Appli
8	606.5	28.1	2732 3 US-08-474-410-60 Sequence 60, Appli
9	606.5	28.1	2732 4 US-08-486-673B-60 Sequence 60, Appli
10	605.5	28.1	1475 1 US-08-097-938-1 Sequence 1, Appli
11	605.5	28.1	1475 1 US-08-476-000-1 Sequence 1, Appli
12	605.5	28.1	1475 1 US-08-472-840-1 Sequence 1, Appli

13	605.5	28.1	1475	2	US-08-476-976-1	Sequence 1, Appli
14	605.5	28.1	1475	3	US-08-474-410-1	Sequence 1, Appli
15	605.5	28.1	1475	4	US-08-486-673B-1	Sequence 1, Appli
16	585.5	27.1	1255	1	US-08-097-938-3	Sequence 3, Appli
17	585.5	27.1	1255	1	US-08-476-000-3	Sequence 3, Appli
18	585.5	27.1	1255	1	US-08-472-840-3	Sequence 3, Appli
19	585.5	27.1	1255	2	US-08-476-976-3	Sequence 3, Appli
20	585.5	27.1	1255	3	US-08-474-410-3	Sequence 3, Appli
21	585.5	27.1	1255	4	US-08-486-673B-3	Sequence 3, Appli
22	585.5	27.1	1255	4	US-08-476-000-62	Sequence 62, Appli
23	572.5	26.5	1414	1	US-08-472-840-62	Sequence 62, Appli
24	572.5	26.5	1414	2	US-08-476-976-62	Sequence 62, Appli
25	572.5	26.5	1414	3	US-08-474-410-62	Sequence 62, Appli
26	572.5	26.5	1414	4	US-08-486-673B-62	Sequence 62, Appli
27	523.5	24.3	4895	3	US-09-053-866-1	Sequence 1, Appli
28	523.5	24.3	4895	4	US-09-479-130-1	Sequence 1, Appli
29	523	24.2	3480	1	US-07-657-769B-68	Sequence 68, Appli
30	523	24.2	3480	1	US-07-789-184-219	Sequence 219, App
31	523	24.2	3480	1	US-08-475-263-219	Sequence 219, App
32	523	24.2	3480	1	US-08-485-886-219	Sequence 219, App
33	523	24.2	3480	2	US-08-477-362-219	Sequence 219, App
34	523	24.2	3480	2	US-08-477-134-219	Sequence 219, App
35	523	24.2	3480	3	US-08-473-489A-219	Sequence 219, App
36	523	24.2	3480	3	US-08-485-695-219	Sequence 219, App
37	523	24.2	3480	4	US-08-018-760-219	Sequence 219, App
38	516.5	23.9	1764	1	US-08-313-553-12	Sequence 12, Appli
39	516.5	23.9	1764	3	US-08-767-993-12	Sequence 12, Appli
40	410.5	19.0	1643	1	US-08-383-750-3	Sequence 3, Appli
41	410.5	19.0	1643	3	US-08-352-678-3	Sequence 3, Appli
42	410.5	19.0	1643	5	PCT-US93-09636-3	Sequence 3, Appli
43	410.5	19.0	1901	1	US-08-153-848-43	Sequence 43, Appli
44	410.5	19.0	1901	3	US-09-299-843A-43	Sequence 43, Appli
45	410.5	19.0	1901	4	US-09-088-337B-43	Sequence 43, Appli

ALIGNMENTS

RESULT 1
US-08-742-440A-4
; Sequence 4, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Fastseq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400

```
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-742-440A-4

Alignment Scores:
Pred. No.: 3,14e-208 Length: 1224
Score: 2142.00 Matches: 405
Percent Similarity: 99.26% Conservatives: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 2 Gaps: 0

US-09-208-629f-6 (1-408) x US-08-742-440A-4 (1-1224)

QY 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleIleYsMet 20
Db 1 TGCCTCCATGATTTCAGAGATTTTCAATACGTTTAAGAGACGGGACTCAGGTCATCAAAATG 60
QY 21 LysAlaIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 61 AAGCCCTCATCTTTGCACGCTGCTGGCCCTCTGCTTCTGTTGCCACATTTTGTGAGAGT 120
QY 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleIleYsThrPheArg 60
Db 121 GGCATGAAATGATACAAACAACTTGGCAAGCAACCTTACCATTAAAGACCTTTGCT 180
QY 61 GlyAlaProPheAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
Db 181 GGAGCTCCCCCAATTTCTTGAAGAGTTCCCTTTTCTGCTTGAAGAGCTTGAAGCTGGACAG 240
QY 81 AlaThrIleThrValIleYsLysCysProGluGluSerAlaSerHisLeuHisValIleYs 100
Db 241 GCCACATTTACTGTAATAATTAAGTCCCTGGAAGAAGTCTTCACTCCATGCCATGTGAAA 300
QY 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
Db 301 AATGCTACCATGGGTACCTGACCACTCTTAAGTACTAACTGATACCTGCCATCTAC 360
QY 121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
Db 361 CTCCTGGTGTGTAGTTGGTGTCCCGGCAATGCTGTGACCTGTGGATGCTTTCTTC 420
QY 141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
Db 421 AGGACCATGATCCATCTGTACCACTGTATTTACACCACTGGCCATTCAGATTTCTT 480
QY 161 PheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlnAsnAsnTrpValPhe 180
Db 481 TTTTGTGTACATTTGCTTTTGAATAGCTTATCTCTCAATGGGAACAACTGGGTATT 540
QY 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 541 GGAGAGGTCTGTGGCGGGCCACCAAGTCATCTTATGGCAACATGTACTGCTCCATT 600
QY 201 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db 601 CTGCTCTTGGCTGCAATCAGCATCAACCGCTACCTGGCCATGTCATCTTTTACCTTAC 660
QY 221 ArgGlyLeuProIleHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 661 CGGGGCTGCCCAAGCACACCTATGCTTGGTAACATGTGGACTGGTGGGGCAACAGTT 720
QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db 721 TTTCTTATATGCTGCCATTTTTCATCTACTGAAGCAGGAATATATCTTCTTGTGAGCAGAC 780
```

```
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db 781 ATCACCACCTGCCATGATGTTCAACACATTTGGAGTCTCTCATCTCCCTTCCAACTCTAT 840
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 841 TACTTCATCTCTTGGCATCTTTGGATTCTTAATTTCCATTTGCTTATCATCTACTATGC 900
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTyrLeuTyrTyrValIys 320
Db 901 TATGAGCCATCATCCGGACACTTAATGCAATACCATCATAGATGTTGTGGTATGTTAAG 960
QY 321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleLeu 340
Db 961 GCGAGTCTCTCATCTTGTGATTTTACCATTTTGTCTTCCAAAGCAATATTATTCTT 1020
QY 341 IleIleHisHisAlaAsnTyrTyrTrpAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
Db 1021 ATTATTACCATGCTACTACTACTACCAACCACTGATGGCTTATATTTATATATCTC 1080
QY 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
Db 1081 ATAGCTTTGTGCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTCTCATG 1140
QY 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys****AsnAspLeuArgGlu 400
Db 1141 TCAAAACCAAGCAATCACTCCACTGCTTACCTTACAAATAGTAGTAATGATCTTAGAGAA 1200
QY 401 GlnGlyGlnProSerGlnArgThr 408
Db 1201 CAAGACAGCCATCACAGAGAAGC 1224

RESULT 2
US-08-742-440A-5
; Sequence 5, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; NUMBER OF SEQUENCES: 3 and Uses Thereof
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1102 base pairs
```

		TYPE: nucleic acid		STRANDEDNESS: double		TOPOLOGY: linear		MOLECULE TYPE: Genomic DNA		US-08-742-440A-5	
		Alignment Scores:		Pred. No.:		Length:		Matches:		Conservative:	
		Score:		1.66e-188		1102		364		1	
		Percent Similarity:		99.46%		1		2		0	
		Best Local Similarity:		99.18%		0		0		0	
		Query Match:		90.26%		2		0		0	
		DB:									
US-09-208-629F-6 (1-408) x US-08-742-440A-5 (1-1102)											
QY	40	SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe	59								
Db	1	ACAGGCATGGAAATGATCAACAACCTTGGCAAGCCCAACCTTACCCATTAGACCTTT	60								
QY	60	ArgGlyAlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr	79								
Db	61	CGTGGAGCTCCCAAAATCTTTGAAGAGTTCCCTTTCTGCTTGGAGGCTGGACA	120								
QY	80	GlyAlaThrIleThrValLysLysCysProGluGluSerAlaSerHisLeuHisVal	99								
Db	121	GGAGCCACGATTAAGTAAATAATTAAGTCCCTGGAAGAAAGTCTTTCACATCTCCATGTG	180								
QY	100	LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrIleLeuLeuProAlaIle	119								
Db	181	AAAAATGCTACCATGGGTACCTGACCACTCTTAAAGTACTAACTGATACCTGCGCATC	240								
QY	120	TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe	139								
Db	241	TACTCTGCTGTGTAGTTGGTGTCCCGGCAATGCTGTGACCTGTGGATGCTTTTC	300								
QY	140	PheArgThrArgSerIleCysThrThrValPheThrAsnLeuAlaIleAlaAspPhe	159								
Db	301	TTCAGGACCAGATCCATCTGTACCACTGTATTCTACCAACCTGGCCATTCAGATTTT	360								
QY	160	LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrpVal	179								
Db	361	CTTTTGTGTACATGTGCTTAAAGATAGCTTATCATCTCAATGGGAACCACTGGGTA	420								
QY	180	PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer	199								
Db	421	TTTGGAGAGTCTGTGCGGGCCACACAGTCATCTTCTATGGCAACATGACTGCTCC	480								
QY	200	IleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr	219								
Db	481	ATTCTGCTCTGCTGTCATCAGCATCAACCGCTACCTGGCCATCGTCCATCTTTCACC	540								
QY	220	TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThr	239								
Db	541	TACCGGGGCTGCGCCCAAGCACACCTATGCTTGGTAACATGTGAGTGTGGGCAACA	600								
QY	240	ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro	259								
Db	601	GTTTCTTATATGCTGCCATTTTTCATCTAGACAGCAATATATCTTGTTCAGCCA	660								
QY	260	AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeu	279								
Db	661	GACATCACCACTGCCATGATGTTCAACACACTTGGAGTCTCTCATCTCCCTCCAACTC	720								
QY	280	TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyr	299								
Db	721	TATTACTTCTATCTCTGTCATCTTGTGGATCTTAAATTCATTTGCTGTATCATCTAC	780								
QY	300	CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpVal	319								
Db	781	TGCTATGAGCCATCATCCGACACCTAATGCAATGATCATGATGTTGTGGTATGTT	840								
QY	320	LysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIle	339								

Alignment Scores:		Pred. No.:		Length:	
Score:		2.45e-137		1224	
Percent Similarity:		1443.50		Matches:	
Best Local Similarity:		83.74%		Conservative:	
Query Match:		71.54%		Mismatch:	
DB:		66.92%		Indels:	
		2		1	
				Gaps:	
				1	

US-09-208-629F-6 (1-408) x US-08-742-440A-1 (1-1224)

QY 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
DB 33 GGGTCTTCAGGACATCAAGATGAATCCCTATCTTGGTTGCGGCTGGCTGCTGTTCTG 92

QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
DB 93 CAGTCACCTGTTGGCAAGTGGCATA---AATGTTTCAGACAACTCAGCAAGGCCAAC 149

QY 54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73
DB 150 TTAACATATTAAAGAGTTTAAATGGGGTCCCAAAATACCTTTGAAGAATTCCCATTTCT 209

QY 74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
DB 210 GACATAGAGGGTGGAGAGGAGCCACCAACTATAAAAGCGGAGTGTCCCGAGGACAT 269

QY 94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTrpLeuThrSerSerLeuSerThr 113
DB 270 ATTTCACTCTCCAGCTGAATATGCTACCATAGGATACCTGAGAGTTCTTTAAGTACC 329

QY 114 LysLeuIleProAlaIleTrpLeuLeuValPheValGlyValProAlaAsnAlaVal 133
DB 330 CAAATGATACCTGCCATCTATATCTGCTGTTTGGTTGGTGTACCATCCAAACATCGTG 389

QY 134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTrpThrAsn 153
DB 390 ACCCTGTGAAATCTCTTTAAGGACCAATCATCATGCTCGTCACTTTCCACACCAAC 449

QY 154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTrpHisLeu 173
DB 450 CTGGCCATCGCAGATCTCTTCTGTGTCACTGCCATTTAAGATCGCCTACCATCTC 509

QY 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTrp 193
DB 510 AATGGCAACAACCTGGGTATTTGGCGAGGTCAATGCGCGATCACCACGGTCTGTTTCTAC 569

QY 194 GlyAsnMetTrpCysSerIleLeuLeuAlaCysIleSerIleAsnArgTrpLeuAla 213
DB 570 GGCACATGATCTGCGCTATCTGATCTCTCATGTCGATGGCATCAACCGCTACCTGGCC 629

QY 214 IleValHisProPheThrTrpArgGlyLeuProLysHisThrTrpAlaLeuValThrCys 233
DB 630 ACGGCTCACCTTTCAATACAGAGCTGCCAAACGCGATCTCTCTGCTCATGTGT 689

QY 234 GlyLeuValTrpAlaThrValPheLeuTrpMetLeuProPhePheIleLysGlnGlu 253
DB 690 GGCATAGTGTGGTCACTGTTTCTTATACATGCTGCCCTTTGTCTCATCTCGAAGCAGGAG 749

QY 254 TyrTrpLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
DB 750 TACCACCTCGTCACCTCAGAGATCACCACCTGCGCAGATGTCTGTCAGCGGTGCGAGTCC 809

QY 274 SerSerProPheGlnLeuTrpTrpPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
DB 810 CATCATCTTCGATCTCATTCTCTCTCTAGCAATCTTTGGGTCTCTCATCCCG 869

QY 294 PheValLeuIleIleTrpCysTrpAlaAlaIleArgThrLeuAsnAlaTrpAspHis 313
DB 870 TTGTGATCATCATCTCTGTGTACAGACTCTCATCCACAACTTAAATCAAAGGATCGG 929

QY 314 ArgTrpLeuTrpTrpValLysAlaSerLeuLeuLeuValIlePheThrIleCysPhe 333
DB 930 ATATGCTGGGTACATCAAGGCGCTCTCTCATCTCTTGTGATTTTCAATTTGCTTT 989

QY 334 AlaProSerAsnIleIleLeuIleHisHisAlaAsnTrpTrpTrpAsnAsnThrAsp 353
DB 990 GCGCCCAACCAACATCATCTGTAATCCACCATGCCAACTACTACTACCAATACCGAC 1049

QY 354 GlyLeuTrpPheIleTrpLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373

DB 1050 AGCTTGTTACTTTATGTATCTTATTGCTGTGCTGGGAGCCTGAATAGCTGCTAGAT 1109

QY 374 ProPheLeuTrpPheLeuMetSerLys 382
DB 1110 CCAATCTTTACTTTTGTCATGTCGAAA 1136

RESULT 4
US-08-742-440A-2
; Sequence 2, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742.440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURES:
US-08-742-440A-2

Alignment Scores:
Pred. No.: 3,22e-88 Length: 1124
Score: 959.00 Matches: 185
Percent Similarity: 73.9% Conservative: 34
Best Local Similarity: 62.50 Mismatches: 53
Query Match: 44.46 Indels: 25
DB: 2 Gaps: 2

US-09-208-629F-6 (1-408) x US-08-742-440A-2 (1-1124)

QY 44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
DB 36 AATGTTTCAGACAACTCAGCAAGGCCAACTTAACTATTAAAGAGTTTAAATGGGGTCCC 95

QY 64 ProAsnSerPheGluGluPhePheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
DB 96 CAAATACCTTTGAAGAATTCNN----- 118

QY 84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103


```
Db 119 .....NNNNNTACAACTCTCCATGTGAATATGCTACC 151
Qy 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeuLeuVal 123
Db 152 ATGGGATACCTGAGAGTTCCTTAAGTACCAAGTGTACCTGCGCATCTACATCTCTGGTG 211
Qy 124 PheValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArg 143
Db 212 TTTGTGATTGGTGTACAGCAACATCGTACCGCTGTGGAACATCTCTCAAGGACCAAA 271
Qy 144 SerLeuLeuThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db 272 TCCATCTGTCTGTCTATCTTTACACCAACCTGGCCATCGCGATCTCTTTCTGTGTC 331
Qy 164 ThrLeuProPheLysIle-AlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluVal 183
Db 332 ACCTGCGGTTTAAGATCANNCTTACCATCTCATGTGCACCAACTGGGTATTGGCGAGGT 391
Qy 183 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer---IleLeuLeu 202
Db 392 CATGTGCGGATCACCACGGTCTTTTCTACGGCAACATGTACTGCGCTANNNTCTGAT 451
Qy 202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPropheThrTyrArgGlu 222
Db 452 CCTCACCCTGCATGGGCATCAACCGCTACCTGGCCACGGCTCACCTTTTCATACACAGAA 511
Qy 222 YLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeu 242
Db 512 GCTGCCCAACGAGCTTCTCCATGTCATGTGTGCGATGGTGTGGGTATGTTCTTCTT 571
Qy 242 uTyrMetLeuProPhePheIle-LeuLysGlnGluTyrTyrLeuValGlnProAspIleT 262
Db 572 ATACATGCTGCCCTTTGTCTATCCNNAAGCAGAGTACCACCTCGTCCACCTCGAGATCA 631
Qy 262 hrThrCysHisAspValHisAsnThrCysGluSerSerSerPropheGlnLeuTyrTyrP 282
Db 632 CCACCTGCCACGATGTGTCACGCGTGGCGANTCCCATCATCTCTTCGATTTCTACTACT 691
Qy 282 heIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrA 302
Db 692 TCGTCTCTTAGCATCTTTGGGTCTCTCATCCGCTTGTGATCATCATCTCTCTGTATCA 751
Qy 302 laIaIleIleArgThrLeuAsnAlaTyrAspHisArgTyrLeuTyrTyrValLysAlas 322
Db 752 CGACTCTCATCCACAACTTAAATCAAAAGATCNGATATGGCTGGGTACATCAAGGCGG 811
Qy 322 erLeuLeuLeuValIlePheThrIleCysPheAlaProSer 336
Db 812 TCCTCTCATCTTGTGAATTCACCATCTGCTTCCCCCCCCACC 855
```

RESULT 5

```
US-08-476-000-60
; Sequence 60, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269
; US-08-476-000-60

Alignment Scores:
Pred. No.: 6,99e-52 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 1 Gaps: 12

US-09-208-629F-6 (1-408) x US-08-476-000-60 (1-2732)
Qy 22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db 85 AGCCTGGGTGGTCTGCTGGGAGGTATCACCTTCTTGGCGGCTCGTCTCTGCGAGCGG 144
Qy 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db 145 ACCGAGAACCCTTGACCCGGGACGCAACACAGTAAGGAAGTCTTTATTGGCAGATTA 204
Qy 60 ArgGlyAlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db 205 GAAACCCACCTCCA-----ATCACT 225
Qy 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db 226 GGGAAAGGGGTTCCGTAGAACACAGGCTTTTCCATCGATGAGTTCTCTGCGTCC----- 279
Qy 98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
Db 280 -----ATCCTCACCGGAAGCTGACCACCGTCTTCTTCTCG 315
Qy 118 AlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db 316 GTGCTGCTATATTGTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
Qy 138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db 376 TTCCCTTTTCCGAACGAAGAAACACACCCCGCGTGATTTTACATGGCCAACTGCGCTTG 435
Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db 436 GCGGACCTCTCTCTGTCTATCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 495
Qy 177 AsnTyrValPheGlyGluValLysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGCTATGTAACATG 555
Qy 197 TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
```

```
Db 556 TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu---- 235
Db 616 CCCATG-----GGACACCCCGAGGAAG---GCAACATCCCGTGGCGTCTCC 663
Qy 236 -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheLeuLeuLysGlnGlu 253
Db 664 TTGGCAATCGGCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACC 723
Qy 254 TyrTyrLeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGluSer 273
Db 724 ATCTACATCCAGCATTTGAACATCACACCTGTCCAGATGTG----- 765
Qy 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
Db 766 ---CTGCTGAGGAGGTATTGGTGGGGACATGTTCAATTACTTCTCTCACTGGCCATT 822
Qy 288 PheGlyPheLeuLeuValPheValLeuIleTyrCysTyrAlaAlaIleIleArgThr 307
Db 823 GGAGTCTTCTGTTCCGGCCCTCTTACTGTCATCGCTACGTCGTCTGATCAAGACG 882
Qy 308 LeuAsnAla-----TyrAspHisArgTrpLeuTyrTyrValLysAla 321
Db 883 CTCGGCTCTCTGCTATGATGAACACTCAGAGAAGAAAGGAGGCGTATCCGACTC 942
Qy 322 SerLeuLeuLeuValPheThrIleCysPheAlaProSerAsnIleLeuLeu 341
Db 943 ATCATCAGCGTGTGCGCCATGATCTCTGCTCTGCTCTAGCAACCTTCTGCTCGTA 1002
Qy 342 IleHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db 1003 GTCATTTTCTTAATCAAAACCCAGAGCAGCAGCGTCTACCCCTCTACCTTGTG 1062
Qy 362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1063 GCCCTCTGCTGTGCGACCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCA 1122
Qy 382 LysThr---ArgAsnHisSer 387
Db 1123 AAAGATTTCAGGATCACGCC 1143

RESULT 6
US-08-472-840-60
; Sequence 60, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269
; US-08-472-840-60

Alignment Scores:
Pred. No.: 6,998-52 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 1 Gaps: 12

US-09-208-629F-6 (1-408) x US-08-472-840-60 (1-2732)
Qy 22 AlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db 85 AGCCTGGGTGGTCTGCTGGAGGTATACCCCTTCTGGGGCGCTCGTCTCTCGACCGCG 144
Qy 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db 145 ACCGAGAACCTTGCACCGGACGCAACCAACAAAGAAAGAGTCTTATTGGCAGATTA 204
Qy 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db 205 GAAACCCAGCTCCA-----ATCCTCACCGGGAAGCTGACCACGCTCTTCTTCGCG 225
Qy 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db 226 GGGAAAGGGGTTCCGGTAGAACAGGCTTTTCCATCGATGAGTTCTCTGCGTCC----- 279
Qy 98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117
Db 280 -----ATCCTCACCGGGAAGCTGACCACGCTCTTCTTCGCG 315
Qy 118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db 316 GTCGCTACATTTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 375
Qy 138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db 376 TTCTCTTTTCCGAACGAAGAAGAAACACCCCGCGTGATTATACATGGCCAAACCTGGCCT 435
Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db 436 GCCACCTCTCTCTGTCATCTGGTTCCCTCGAAGATCTCTACACCTACATGCGCAAC 495
Qy 177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 496 AACTGGGTCTACGGGGAGCCCTGTGCAAGGTGCTCATTTGGGCTTTTCTATGGTAACATG 555
Qy 197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db 556 TATTGCTCATCTCTTCTATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 615
Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db 616 CCCATG-----GGACACCCCGAGGAAG---GCAACATCCCGTGGCGTCTCC 663
Qy 236 -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
```

```
Db 664 TTGGCAATCGGCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCATGAGCAGACC 723
QY 254 TTTTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db 724 ATCTACATCCAGCATTTGACATCACCACTGTACAGATGTG-----765
QY 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
Db 766 ---CTCGCTGAGGAGGATTTGGTGGGGACATGTTCAATTACTTCTCTCACTGGCAATT 822
QY 288 PheGlyPheLeuIleProPheValLeuLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db 823 GAGTCTTCTGTTCCCGGCCCTCTTACTGTCATCTGCTACGTGCTCATGATCAAGACG 882
QY 308 LeuAsnAla-----TyrAspHisArgTyrPheLeuTyrValLysAla 321
Db 883 CTCGGCTCTCTGCTATGATGATGACACTCAGAGAAAGGAGGAGGCTATCCGACTC 942
QY 322 SerLeuLeuLeuValPhePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db 943 ATCATCAGCGTGTGGCCATGTTACTTCTGCTTCTAGCAACCTTCTGCTCGTA 1002
QY 342 IleHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db 1003 GTGCATTATTTCCTAATCAAAACCCAGAGCAGCCAGCTCTACGCCCTCTACCTTGT 1062
QY 362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1063 GCCTCTGCTGTCGACCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCA 1122
QY 382 LysThr---ArgAsnHisSer 387
Db 1123 AAAGATTTCAGGATCAGGCC 1143
```

RESULT 7

```
US-08-476-976-60
; Sequence 60, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
```

SEQUENCE CHARACTERISTICS:

```
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269
US-08-476-976-60
```

Alignment Scores:

```
Pred. No.: 6,998-52 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 2 Gaps: 12
```

US-09-208-629F-6 (1-408) x US-08-476-976-60 (1-2732)

```
QY 22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db 85 AGCCTGGCGTGTCTGCTGGAGGTATCACCTTCTGGCGGCTCGGTCTCTCGACCCGG 144
QY 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db 145 ACCGAGAACCTTGACCGGGACGCAACACAGTAAAGGAAGAGTCTTATTGGCAGATT 204
QY 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db 205 GAAACCCAGCCCTCCA-----ATCACT 225
QY 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db 226 GGGAAAGGGTTCGGTAGAACACGAGCTTTTCATCGATGAGTTCTCTCGCTCC----- 279
QY 98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuPro 117
Db 280 -----ATCCTCACCGGGAAGCTGACCACGGTCTTTCTTCGCG 315
QY 118 AlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db 316 GTCTGTCTACATTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 375
QY 138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db 376 TTCTTTTCCGAAACGAAAGAACACCCCGCGGTGATTATCATGGCCAACTTGGCCTTG 435
QY 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db 436 GCGGACCTCTCTCTGTCTATCTGGTTCCGCTGAAAGATCTCTACCACTACATGGCAAC 495
QY 177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 496 AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTATGATG 555
QY 197 TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db 556 TATTGCTCCATCTCTTCTATGACCTGCTCAGCGGTGAGAGGTACTGGGTGATCGTGAAC 615
QY 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db 616 CCCATG-----GGACACCCCGAGGAAG---GCAACACATCGCGCTTGGCGTCTCC 663
QY 236 -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db 664 TTGCAATCTGGCTCTCTGATTTTCTGTGTACCATCCCTTTGTGTGTGTGTGTGTGTGTGT 723
QY 254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db 724 ATCTACATTCCAGCATTTGAACATCACCACTCTGTCAGATGTG-----765
QY 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
```

```

Db      766  ---CTGCTGAGGAGTATTGGTGGGGACATGTTCAATTACTTCTCTCACTGGCCATT 822
Qy      288  PheGlyPheLeuLeuPheValLeuLeuLeuTyrCysTyrAlaAlaLeuLeuArgThr 307
Db      823  GGAGTCTTCTGTTCCCGCCCTCTTACTGTCATCGCTACGTCGTCTGATCAAGACG 882
Qy      308  LeuAsnAla-----TyrAspHisArgTyrPheLeuTyrValLysAla 321
Db      883  CTCGGCTCTCTGCTATGATGAACACTCAGAACGAAGAGGAGGAGGCTATCCGACTC 942
Qy      322  SerLeuLeuLeuValLeuPheThrLeuPheAlaProSerAsnLeuLeuLeuLeu 341
Db      943  ATCATCAGCTGCTGGCCATGATCTTCTGCTTCTCTAGCAACCTTCTGCTCGTA 1002
Qy      342  IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheLeuTyrLeuLeu 361
Db      1003  GTCCATTTTTCCTAATCAAAACCCAGAGCAGACGCTACGCCCTCTACCTTGTG 1062
Qy      362  AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPheLeuTyrPheLeuMetSer 381
Db      1063  GCCTCTGCTGTCGACCTCAACAGCTGATAGACCCCTTGTCTATTACTTTGTCTCA 1122
Qy      382  LysThr---ArgAsnHisSer 387
Db      1123  AAAGATTTCAGGATCAGGCC 1143

RESULT 8
US-08-474-410-60
; Sequence 60, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,410
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 60:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2732 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 73..1269

```

```

US-08-474-410-60
Alignment Scores:
Pred. No.: 6,998-52 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 3 Gaps: 12

US-09-208-629F-6 (1-408) x US-08-474-410-60 (1-2732)
Qy      22  AlaLeuLeuPheAlaAlaGlyLeuLeuLeuLeuLeuPheCysGlnSerGly 41
Db      85  AGCCTGGCGTGGCTCTGGAGGTATCACCCTTCTGGCGCCCTCGTCTCTCGACCGG 144
Qy      42  MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db      145  ACCGAGAACCTTGCACCGGACGCAACACAGTAAGGAAGAGTCTTATTGGCAGATTA 204
Qy      60  ArgGlyAlaProProAsnSerPheGluGluPhePheSerAlaLeuGluGlyTrpThr 79
Db      205  GAAACCCAGCCCTCCA-----ATCACT 225
Qy      80  GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db      226  GGGAAAGGGTTCGGTAGAACACGCTTTCATCGATGAGTCTCTCGCTCC----- 279
Qy      98  HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuPro 117
Db      280  -----ATCCTCACCGGAAGCTGACCACGGTCTTCTTCTCG 315
Qy      118  AlaIleTyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db      316  GTCGCTCATATTGTTGTGTTGTTGTTGCCAGTAAGTGGCCCTCTGGATG 375
Qy      138  LeuPhePheArgThrArgSerIleCysThrValPheTyr---ThrAsnLeuAlaIle 156
Db      376  TTCCCTTTTCCGAACGAAGAAACACCCCGCGTGATTATACATGGCCACCTGGCCTTG 435
Qy      157  AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db      436  GCCGACCTCTCTCTGTCATCTGTTCCCTCGAAGATCTCTACACCTACATGSCAAC 495
Qy      177  AsnTrpValPheGlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMet 196
Db      496  AACTGGGTCTACGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATG 555
Qy      197  TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db      556  TATTGCTCATCTCTTTCATGACCTGCTCAGCGTGACAGGTACTGGGTGATCGTGAAC 615
Qy      217  PropheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db      616  CCCATG-----GGACACCCCGAAGAAAG---GCAAAACATCGCCGCTGCTCC 663
Qy      236  -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db      664  TTGGCAATCTGGCTCTGCTGTTTCTGTGTCACCATCTTGTATGTCAATGAAGACAGC 723
Qy      254  TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db      724  ATCTACATTCCAGCATTGAACATCACCACCTGTCTCAGTGTG----- 765
Qy      274  SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
Db      766  ---CTGCTGAGGAGGTATTGGTGGGGACATGTTCAATTACTTCTCTCACTGGCCATT 822
Qy      288  PheGlyPheLeuLeuProPheValLeuIleIleTyrCysTyrAlaAlaIleLeuTrp 307
Db      823  GGAGTCTTCTGTTCCCGCCCTCTTACTGCTATCGCTGCTCATGATCAAGACG 882
Qy      308  LeuAsnAla-----TyrAspHisArgTyrPheLeuTrpLeuTyrValLysAla 321

```


CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 CITY: Washington, D.C.
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/097,938
 FILING DATE: 26-JUL-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: MURASHIGE, KATE H.
 REGISTRATION NUMBER: 29,959
 REFERENCE/DOCKET NUMBER: 22803-20006.00
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1475 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 232..1416
 FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 232
 US-08-097-938-1

Alignment Scores:
 Pred. No.: 3,616-52 Length: 1475
 Score: 605.50 Matches: 140
 Percent Similarity: 52.50% Conservative: 70
 Best Local Similarity: 35.00% Mismatches: 117
 Query Match: 28.07% Indels: 74
 DB: 1 Gaps: 13

US-09-208-629F-6 (1-408) x US-08-097-938-1 (1-1475)

QY 18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
 Db 221 GTTCATCTGAATGTTCCATTAAACACAGCAGCCTTACTGTTGGACCATTTATCTCAG 280
 QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
 Db 281 TAATGATTTCTGCTCCGCTTTCTTTGTACAGGAGCGCAACACAGTAAAGGAAGTCTTA 340
 QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
 Db 341 TTGGCAGATTAGAACCCACCTCCATCACTCGGAAGGGGTTCCGGTA-GAACAGGC 399
 QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
 Db 400 TTTTCCATCATGAGTTC----- 417
 QY 85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAlaThrMet 104
 Db 418 -----TCTCGCTCC----- 426
 QY 105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
 Db 427 ---ATCCTCACCGGGAAGCTGACACCGCTTTCTTCGGCTGCTACATTATGTTT 483
 QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
 Db 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144

Db 484 GTGATTGGTTTGGCCAGTAATGGCATGGCCCTCTCGATCTTCTCTTTCGACGAGAGAG 543
 QY 145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
 Db 544 AACACACCGCGCGTATTACATGGCCAACTTGGCGCTTGGCCGACCTCTCTCTGTCTCATC 603
 QY 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
 Db 604 TGGTTCCCGCTCGAAGATCTCTACCACTACATGCAACAACCTGGTCTACGGGAGGCC 663
 QY 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
 Db 664 CTGTGCAAGGTGCTCATTTCTTCTATGTAACATGATTTGCTCCATCTCTCTCATG 723
 QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
 Db 724 ACCTGCCTCAGCGTCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
 QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
 Db 775 CCCAGGAAGAAG--GCAAAACATCGCGTGGCGCTCTCTTGGCAATCTGGCTCTGATT 831
 QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
 Db 832 TTTCTGTGTCACCATCTCTTGTATGTCATGAAGCAGACCATCTACATTCCAGCATTTGAAC 891
 QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
 Db 892 ATCACCACTGTCAGATGTG-----CTGCCTGAGGAGGTATTG 930
 QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
 Db 931 GTGGGGACATGTTCAATTACTTCTCTCAGTGGCATTTGGAGTCTTCTGTTCGGGCC 990
 QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
 Db 991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGAGCTCCGCTTCTCTGATGGAT 1050
 QY 311 -----TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuValIle 328
 Db 1051 GAACACTCAGAGAAACAAAGGAGGCTATCCGACTCATCATCCGTGCTGGCCATG 1110
 QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
 Db 1111 TACTTCACTGCTTTGCTCTTAGCAACCTCTGCTCGTAGTCATATTTCCTAATCAA 1170
 QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
 Db 1171 ACCCAGAGCAGAGCCAGCTCTACGCCCTCTACCTTGTGCGCCTCTGCTGCGACCTC 1230
 QY 369 AsnSerCysLeuAspPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
 Db 1231 AACAGCTGATAGACCCCTTTGCTATTACTTTGTTCTCAAAAGATTTCAGGATACGCC 1290

RESULT 11

US-08-476-000-1
 Sequence 1, Application US/08476000
 Patent No. 5716789
 GENERAL INFORMATION:
 APPLICANT: SUNDELIN, JOHAN
 APPLICANT: SCARBOROUGH, ROBERT M.
 TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 NUMBER OF SEQUENCES: 63
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORRISON & FOERSTER
 STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20006-1812
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/476,000
 FILING DATE: 07-JUN-1995

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/390,301
 FILING DATE: 25-JAN-1995

ATTORNEY/AGENT INFORMATION:
 NAME: ADLER, REID G.
 REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 2803-0006.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763

TELEX: 90-4030
 INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 1475 base pairs
 TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:
 NAME/KEY: CDS
 LOCATION: 232..1416

FEATURE:
 NAME/KEY: mat_peptide
 LOCATION: 232

US-08-476-000-1

Alignment Scores:

Pred. No.:	3 61e-52	Length:	1475
Score:	605.50	Matches:	140
Percent Similarity:	52.50%	Conservative:	70
Best Local Similarity:	35.00%	Mismatches:	117
Query Match:	28.07%	Indels:	74
DB:	1	Gaps:	13

US-09-208-629f-6 (1-408) x US-08-476-000-1 (1-1475)

```

QY 18 IleLysMetLysAlaLeuLeuPheAlaAlaGlyLeuLeuLeuLeuProThr--- 36
   :::::
Db 221 GTTCATCTGAATGTTCCATTAAACACAGCAGCCTTACTGTGGACCATTTATCTCAG 280
   :::::
QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
   :::::
Db 281 TAATGATTCTGCTCCCGCTTTCTTTGTACAGGACGCCAACACAGGAAGAGTCTTA 340
   :::::
QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
   :::::
Db 341 TTGGCAGATTAGAAACCCAGCCTCCATCTACTGGGAAGGGTTCGGTA-GAACACAGGC 399
   :::::
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
   :::::
Db 400 TTTTCCATCATCAGTTC----- 417
   :::::
QY 85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
   :::::
Db 418 -----TTGCGTCC----- 426
   :::::
QY 105 GlyTyThrLeuThrSerSerLeuThrLysLeuLeuProAlaIleTyThrLeuLeuValPhe 124
   :::::
Db 427 ---ATCCTACCGGGAAGCTGACACGGCTTCTTCGGGTGCTACATATTGTGTTT 483
   :::::
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
   :::::
Db 484 GTGATTGGTTGCCAGTAATGTCATGGCCCTCTGGATCTCTCTTCGACGAAGAAG 543
   :::::
QY 145 IleCysThrThrValPheTyThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
   :::::
Db 544 AAACACCCCGCGGTGATTATACATGGCCAACTGGCGCTTGGCCGACCTCTCTCTGTCATC 603
   :::::

```

```

QY 164 ThrLeuProPheLysIleAlaTyThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
   :::::
Db 604 TGGTCCCTCGAAGATCTCTACACCTTACACCACTGGCACTGGTCTACGGGAGGCC 663
   :::::
QY 184 LeuCysArgAlaThrThrValIlePheTyThrGlyAsnMetTyCysSerIleLeuLeu 203
   :::::
Db 664 CTGTGCAAGGTGCTCATTTGGCTTTTCTATGTGTAAACATGATTGTCTCCATCTCTTCATG 723
   :::::
QY 204 AlaCysIleSerIleAsnArgTyThrLeuAlaIleValHisProPheThrTyArgGlyLeu 223
   :::::
Db 724 ACCTGCTCAGGTGCGAGGACTGGGTGATCGTGAACCCCATG-----GGACAC 774
   :::::
QY 224 ProLysHisThrTyAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
   :::::
Db 775 CCAGGAAGAAG--GCAACATCGCGGTGGCGTCTCTTGGCAATCTGGCTCTGATT 831
   :::::
QY 241 PheLeuTyMetLeuProPhePheIleLeuLysGlnGluTyThrLeuValGlnProAsp 260
   :::::
Db 832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGACGACCATCTACATCCAGCATTTGAAC 891
   :::::
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyThr 280
   :::::
Db 892 ATCACCACTGTACAGTGT-----CTGCCTGAGGAGGTATTG 930
   :::::
QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
   :::::
Db 931 GTGGGGACATGTTCAATTACTTCTCTCAGTGGCATTTGGAGTCTTCTGTTCGCGCC 990
   :::::
QY 295 ValLeuIleIleTyThrCysTyAlaAlaIleIleArgThrLeuAsnAla----- 310
   :::::
Db 991 CTCCTTACTGTCATCTGCTTACGTACGTGTCATGATCAAGACGCTCGCTCTCTCTGTATGAT 1050
   :::::
QY 311 -----TyrAspHisArgTrpLeuTrpTyThrValLysAlaSerLeuLeuValIle 328
   :::::
Db 1051 GAACACTCAGAAACAAAGGAGGAGGCTATCCGACTCATCATCCGTGTGGCCATG 1110
   :::::
QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyThr 348
   :::::
Db 1111 TACTTCATCTGCTTGTCTCTAGCAACCTTCTGCTCGTAGTGCATTTATTCCTAATCAA 1170
   :::::
QY 349 TyrAsnAsnThrAspGlyLeuTyThrPheIleTyThrLeuIleAlaLeuCysLeuGlySerLeu 368
   :::::
Db 1171 ACCCAGAGGACAGGACGACGCTACGCCCTCTACCTTGTGCGCTCTGCTGTGCGACCTC 1230
   :::::
QY 369 AsnSerCysLeuAspPheLeuTyThrPheLeuMetSerLysThr---ArgAsnHisSer 387
   :::::
Db 1231 AACAGCTGCATAGACCCCTTGTCTATTACTTTGTCTCTCAAAAGATTTACAGGATCACGCC 1290
   :::::

```

RESULT 12

US-08-472-840-1
 ; Sequence 1, Application US/08472840
 ; Patent No. 5763575
 ; GENERAL INFORMATION:

APPLICANT: SUNDELIN, JOHAN
 APPLICANT: SCARBOROUGH, ROBERT M.

TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR

NUMBER OF SEQUENCES: 63

CORRESPONDENCE ADDRESS:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,840

FILING DATE: 435
CLASSIFICATION: 435
PRIOR APPLICATION DATA: US/08/390,301
APPLICATION NUMBER: 2803-0006.20
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1475 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 232..1416
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 232
US-08-472-840-1

Alignment Scores:
Pred. No.: 3,61e-52 Length: 1475
Score: 605.50 Matches: 140
Percent Similarity: 52.50% Conservative: 70
Best Local Similarity: 35.00% Mismatches: 117
Query Match: 28.07% Indels: 74
DB: 1 Gaps: 13

US-09-208-629f-6 (1-408) x US-08-472-840-1 (1-1475)

184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeu 203
664 CTGTGCAAGGTGCTCATTTGGCTTTTCTATGTAACATGATTTGCTCCATCTCTCATG 723
204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
724 ACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTAlaThrVal 240
775 CCCGAGAAG-----GCAACATCCCGTGGCTCTCTGGCAATCTGGCTCTCATG 831
241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
832 TTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGACCATCTACATCCAGCATGAAC 891
261 IleThrThrCysHisAspValHisThrCysGluSerSerSerProPheGlnLeuTyr 280
892 ATCACCACCTGTACGATGTG-----CTGCCTGAGGAGGTATTG 930
281 -----TyrPheIleSerLeuAlaPhePheGlyPheIleProPhe 294
931 GTGGGGGACATGTTCAATTACTTCTCTCACTGGCCATGGAGTCTTCTGTTCGGGCC 990
295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTCTCTATGGAT 1050
311 -----TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuValLeu 328
1051 GAACACTCAGAGAACAAAGGAGGAGGCTATCCGACTCATCATCCGCTGTGCGCCATG 1110
329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
1111 TACTTCATCTGCTTTGCTCCTAGCAACCTTCTGCTGAGTGATATTTCCTAATCAAA 1170
349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
1171 ACCCAGAGCAGAGCCAGCTCTACGCCCTCTACCTTTGCGCCCTCTGCTGTGCGCCCTC 1230
369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
1231 AACAGCTGCATAGACCCCTTTGCTTACTATTGTTCTCAAAAGATTTTCAGGGATCAGCC 1290

RESULT 13
US-08-476-976-1
Sequence 1, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSER: MORRISON & FORSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,976
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: ADLER, REID G.
 REGISTRATION NUMBER: 30,988
 REFERENCE/DOCKET NUMBER: 2803-0006.20
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 887-1500
 TELEFAX: (202) 887-0763
 TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1475 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS
 LOCATION: 232..1416
 FEATURE: mat_peptide
 NAME/KEY: 232
 LOCATION: 232

US-08-476-976-1

Alignment Scores:

Pred. No.: 3,61e-52 Length: 1475
 Score: 605.50 Matches: 140
 Percent Similarity: 52.50% Conservative: 70
 Best Local Similarity: 35.00% Mismatches: 117
 Query Match: 28.07% Indels: 74
 DB: 2 Gaps: 13

US-09-208-629f-6 (1-408) x US-08-476-976-1 (1-1475)

QY 18 IleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuProThr--- 36
 DB 221 GTTCATCTGAAGTTCCTTAAACACACAGCAGCTTACTGTGGACCATTTATCTCAG 280
 QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
 DB 281 TAATGATTCCTGCTCCGCTTTCTTTGTACAGGAGCAACACACAGTAAAGGAAGTCTTA 340
 QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
 DB 341 TTGGCAGATTAGAAACCCAGCTTCAATCACTCGGAAAGGGTTCGGTA-GAACAGGC 399
 QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyThrGlyAlaThrIleThr 84
 DB 400 TTTTCCATCATGAGTTC----- 417
 QY 85 ValLysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
 DB 418 -----TCTCGCTCC----- 426
 QY 105 GlyThrLeuThrSerSerLeuThrLysLeuIleProAlaIleThrLeuLeuValPhe 124
 DB 427 ---ATCTCACCGGGAAGCTGACACGGCTTTCTTCCGGTCTGTACATATTGTGTT 483
 QY 125 ValValGlyValProAlaAsnAlaValThrLeuThrMetLeuPheArgThrArgSer 144
 DB 484 GTGATTGGTTGGCCAGTATGTCATGGCCCTCTGGATCTCTTTCCGAAGAGAAG 543
 QY 145 IleCysThrThrValPheThr-----ThrAsnLeuAlaIleAlaAspPheCysVal 163
 DB 544 AAACACCCCGCTGATTTACATGGGCAACCTGGCTTGGCCGACCTCTCTCTGTCATC 603
 QY 164 ThrLeuProPheLysIleAlaThrHisLeuAsnGlyAsnAsnThrValPheGlyGluVal 183
 DB 604 TGGTCCCCCTGAAGATCTCTACCACTACATGGCAACAACTGGGTCTACGGGGAGGCC 663
 QY 184 LeuCysArgAlaThrThrValIlePheThrGlyAsnMetThrCysSerIleLeuLeuLeu 203
 DB 664 CTGTGCAAGGTGTCATTTGCTTTCTATGTTAAACATGATTGCTCCATCTCTCATG 723
 QY 204 AlaCysIleSerIleAsnArgThrLeuAlaIleValHisProPheThrThrArgGlyLeu 223

DB 724 ACCTGCTCAGCGTCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
 QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
 DB 775 CCCAGGAAGAAG--GCAACATCCCGCTTGGCGTCTCTTGGCAATCTGGCTCTGATT 831
 QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
 DB 832 TTTCTGGTCACCATCCCTTTGATGTATGATGAAGACAGCATCTACATCCAGCATGAAC 891
 QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
 DB 892 ATCACACCTGTACGATGTG-----CTGCCTGAGGAGGTATTG 930
 QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
 DB 931 GTGGGGGACATGTTCAATTACTTCTCTCCTGAGCATGGAGTCTTCTGTTCGCGGC 990
 QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
 DB 991 CTCCTTACTGTCATCTGCTTACGTGCTCATGATCAAGACGCTCGCTCTTCTGCTATGAT 1050
 QY 311 -----TyrAspHisArgTrpLeuThrTyrValLysAlaSerLeuLeuLeuValIle 328
 DB 1051 GAACACTCAGAAACAAAGGAGGAGGCTATCCGACTCATCATCCGTGTGGCCATG 1110
 QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
 DB 1111 TACTTCATCTGCTTTGCTCTAGCAACCTTCTGCTGAGTGCATATTTCCTAATCAAA 1170
 QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
 DB 1171 ACCCAGAGCAGACGACGCTACGCCCTCTACCTTGTGCGCTCTGCTGCGACCTC 1230
 QY 369 AsnSerCysLeuAspPheLeuTyrPheLeuMetSerLysThr-----ArgAsnHisSer 397
 DB 1231 AACAGCTGCATAGACCCCTTCTGCTATTACTTTGTCTCAAAAGATTTTCAGGATCACGCC 1290

RESULT 14
 US-08-474-410-1
 ; Sequence 1, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELIN, JOHAN
 ; APPLICANT: SCARBOROUGH, ROBERT M.
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:

Qy 18 IleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeuProThr--- 36

Db	221	GTTCATCTGAAATGTCATTTAAACACACAGCAGCCTTACTGTGGACCAATTTATCTCAG	280
Qy	37	-----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla	50
Db	281	TAATGATTCGCTCCGCTTCTTTGTACAGGACGCAACAACAGTAAGGAAGTCTT	340
Qy	51	-----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro	64
Db	341	TTGGCAGATTAGAAACCCAGCCTCAATCACTGGGAAAGGGTTCGGTA-GAACCAAGC	399
Qy	65	AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr	84
Db	400	TTTTCATCGATGAGTTC-----TCTCGCTCC-----	417
Qy	85	ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet	104
Db	418	-----TCTCGCTCC-----	426
Qy	105	GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe	124
Db	427	---ATCCTCACCGGAAGCTGACCACGGTCTTCTTCGGTCGCTACATATTATGTTT	483
Qy	125	ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer	144
Db	484	GTGATTGGTTGGCCAGTAATGGCATGGCCCTCTGGATCTTCTCTTCCTCCGAACGAAGAAG	543
Qy	145	IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal	163
Db	544	AAACACCCCGCGTGATTTACATGGCCAACTGGCCCTTGGCCGACCTCTCTCTGTCTCATC	603
Qy	164	ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal	183
Db	604	TGGTTCCCTCGAAGATCTCTACCACTACATGGCAACAACCTGGGTCTACGGGGAGGCC	663
Qy	184	LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu	203
Db	664	CTGTGCAAGGTGCTCATTTGGCTTTTCTATGGTAACATGATTGCTCCATCTCTTCATG	723
Qy	204	AlaCysIleSerIleAsnAtgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu	223
Db	724	ACCTGCCTCAGCGTGACAGGTACTGGGTGATCTGTGAACCCCATG-----GGACAC	774
Qy	224	ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal	240
Db	775	CCAGGAAGAAG---GCAAACTACCGCTGGCGCTCTCTTGGCAATCTGGCTCTGATT	831
Qy	241	PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp	260
Db	832	TTTCTGTGTACCAATCCCTTTGTATGTATGAAGCAGACCATCTACATTCACGACATTGAAC	891
Qy	261	IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr	280
Db	892	ATCACCACTGTTCACGATGTG-----CTGCTCAGGAGGATATG	930
Qy	281	-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe	294
Db	931	GTGGGGGACATGTTCAATTACTTCTCTCACTGGCCATGGAGTCTTCTGTTCCCGGCC	990
Qy	295	ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----	310
Db	991	CTCCTTACTGCATCTGGCTACGTCTCATGATCAAGACGCTCCGCTCTTCTGTGATGGAT	1050
Qy	311	-----TyrAspHisArgTrpLeuTyrTrpValLysAlaSerLeuLeuLeuValIle	328
Db	1051	GAACACTCAGAGAACAAAGACGAGGGCTATCCGACTCATCATCAGCCGTGGCCCAT	1110
Qy	329	PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyr-Tyr	348
Db	1111	TACTTCATCTGCTTTGTCTCTAGCAACTTCTGCTGTAGTCGATTATTCTCTAATCAA	1170
Qy	349	TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu	368

```
Db 302 TACACAAATTGTGTTTGTGGTGGTTTCCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 361
Qy 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCGGAACCTAAGAAAGACCCCTGCTGTGATTACATGGCAATCTGGCCCTTGGCTGAC 421
Qy 159 PheLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db 422 CTCCTCTCTCTCATCTGTTGCCCTTGAAGATTGCTATACATACATGATGCAACAACTGG 481
Qy 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 482 ATTTATGGGAAGCTCTTTGTAATGTCTATTGGCTTTTCTATGGCAACATGACTGT 541
Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 542 TCATCTCTCTATGACCTGCTGAGTGGTGTGAGAGGATTTGGGTATCTGTGAACCCCATG 601
Qy 219 ThrTyrArgGlyLeuProIleHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 602 GGSCAC---TCCAGGAAGAAAGCAACATTTGCCATTGGCATCTCCCTGGCAATATGGCTG 658
Qy 239 ThrValPheLeuTyrMetLeuProPheIleLeuGlyGlnGluTyrTyrLeuValGln 258
Db 659 CTGATCTGCTGCTACCATCCCTTTGTATGTGTGTAAGCAGACCATCTTCATTCCTGCC 718
Qy 259 ProAspIleThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 719 CTGAACATCAGCACCTGTCTATGATGTT-----TTGCCTGAGCAG 757
Qy 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db 758 CTCTTGGTGGGACATGTTCAATTACTTCCCTCTCTGCGCATTTGGGTCTTTCTGTTC 817
Qy 293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
Db 818 CCAGCCTTCTCAGACCTCTGCTATGTGCTGATGATGATGATGATGATGATGATGATGATG 877
Qy 311 -----TyrAspHisArgTrpLeuTyrTyrValIleAlaSerLeuLeuIleLeu 326
Db 878 ATGGATGAACCTCAGAGAAAGAAAGAGAGAGGCGCATCAAACTCATTTGCTACTGTCCTG 937
Qy 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn 346
Db 938 GCCATGCTGATCTGCTTCACTCTAGTAACCTTCTGCTGTGGTGAT----- 988
Qy 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 989 TATTTTCTGATTAAGAGCCAGGCGCCAGAGCCATGCTATGCCCCTGTACATTTAGCCCTC 1048
Qy 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 1049 TGCCTCTTACCTTAAACAGTGCATCGACCCCTTTGCTATTACTTTTGTTCACATGAT 1108
Qy 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGAGCATGCAAGAAAGCTCTCTTTGCCGAGTGTCCGCACTGTAAAGCAG 1165

RESULT 17
US-08-476-000-3
; Sequence 3, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
```

```
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
; US-08-476-000-3

Alignment Scores:
Pred. No.: 3,06e-50 Length: 1255
Score: 585.50 Matches: 124
Percent Similarity: 54.8% Conservative: 62
Best Local Similarity: 36.5% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 1 Gaps: 10

US-09-208-629F-6 (1-408) x US-08-476-000-3 (1-1255)

Qy 80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db 212 GGAGTTACAGTTTGAACACAGTCTTTTCTGTGGATGAGTTTCTGCACT----- 259
Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db 260 -----GTCTCTCACTGGAAACCTGACCACTGCTCTTCTTCCAATTGTC 301
Qy 120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 302 TACACAAATTGTGTTTGTGGTGGTTTCCCAAGTAACGGCATGGCCCTGTGGGTCTTTCT 361
Qy 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCGGAACCTAAGAAAGACCCCTGCTGTGATTATGATGCGCAATCTGGCCCTTGGCTGAC 421
Qy 159 PheLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db 422 CTCCTCTCTCTCATCTGTTGCCCTTGAAGATTGCTATACATACATGATGCAACAACTGG 481
Qy 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 482 ATTTATGGGAAGCTCTTTGTAATGTCTATTGGCTTTTCTATGGCAACATGACTGT 541
Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 542 TCATCTCTCTATGACCTGCTGAGTGGTGTGAGAGGATTTGGGTATCTGTGAACCCCATG 601
```

QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrrpAla 238
Db 602 GGCAC---CCAGGAAGAGCAACATTGGCATCGCATCCCTGGCAATATGGCTG 658
QY 239 ThrValPheLeuTyrMetLeuProPhePheIleuLysGlnGluTyrTyrLeuValGln 258
Db 659 CTGATTCGTGCTGCTACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATCTCTGCC 718
QY 259 ProAspIleThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 719 CTGAACATCAGACCTGTCATGATGTT-----TTGCCTGAGCAG 757
QY 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuLeu 292
Db 758 CTCTGTGGGAGACATGTTCAATTACTCTCTCTGCGCCATTGGGGCTTTCTGTTTC 817
QY 293 ProPheValLeuLeuIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
Db 818 CCAGCTTCCTACAGCCTCTGCTATGCTGCTGATGATGATGCTGCTGATCTCTGCC 877
QY 311 -----TyrAspHisArgTrrpLeuTrrpTyrValLysAlaSerLeuLeuLeu 326
Db 878 ATGATGAAACATCAGAGAGAAAGAGAGAGGCGCATCAAACTCATTTGCTGCTGCTG 937
QY 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn 346
Db 938 GCCATGTACCTGATCTGCTTCACTCTCTAGTAACCTTCTGCTGCTGCTGCTGCTG 988
QY 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuLeuAlaLeu 363
Db 989 TATTTTCTGATTAGAGCCAGGCGCCAGAGCCATGCTATGCTGCTGCTGCTGCTGCTG 1048
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspPheLeuTyrPheLeuMetSerLysThr 383
Db 1049 TGCTCTCTACCTTACAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1108
QY 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGATCATGCAAGAACGCTCTCTCTTGGCGAAGTGTCCGCACTGCTGAAAGCAG 1165

RESULT 18
US-08-472-840-3
; Sequence 3, Application US/08472840
; Patent No. 5763575
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/472,840
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988

REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
; US-08-472-840-3
Alignment Scores:
Pred. No.: 3,06e-50 Length: 1255
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 1 Gaps: 10
US-09-208-629f-6 (1-408) x US-08-472-840-3 (1-1255)
QY 80 GlyAlaThrIleThrValLysIleLysCysProGluSerAlaSerHisLeuHisVal 99
Db 212 GGAGTTACAGTTGAAACAGTCTTTCTGTGGATGAGTTTCTGTCATCT----- 259
QY 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIle 119.
Db 260 -----GTCCTCACTGGAAACTGACCACCTGCTCTCTTCCATGCTCAATGTC 301
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrrpMetLeuPhe 139
Db 302 TACACATTTGTTTGTGGTGGTTTGCCAGTACGGCATGGCCCTGTGGGTCTTTCTT 361
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCCGAACCTAAGAAGACACCTGCTGTGATTACATGCGCAATCTGCGCTTGGCTGAC 421
QY 159 PheLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAsnTrrp 178
Db 422 CTCTCTCTGTCATCTGTTCCCTTGAAGATTGCTATCATCATATGCGCAACACTGG 481
QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTrrpCys 198
Db 482 ATTTATGGGAAGCTTTTGTATGCTTATGCTTATGCTTTTCTATGGCAACATGTA 541
QY 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 542 TCCATTCTCTCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrrpAla 238
Db 602 GGCAC---TCCAGGAAGAGCAACATTGGCATCTCTCTGCGCAATATGGCTG 658
QY 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 659 CTGATTCGTGCTGCTACCATCCCTTTGTATGTCGTGAAGCAGACCATCTTCATCTGCC 718
QY 259 ProAspIleThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 719 CTGAACATCAGACCTGTCATGATGTT-----TTGCCTGAGCAG 757
QY 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuLeu 292
Db 758 CTCTGTGGGAGACATGTTCAATTACTCTCTCTGCGCCATTGGGGCTTTCTGTTTC 817

Db 1049 TGCCCTCTACCCCTTAACAGCTGATCGACCCCTTTGTCTATTACTTTGTTCACATGAT 1108
Qy 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGATCATGCAAGAAGCGCTCTCCITTCGCGAAGTGTCGCCGACGTGTAAGCAG 1165

RESULT 20

US-08-474-410-3
; Sequence 3, Application US/08474410
; Patent No. 6043212
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/474,410
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 56..1249
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 56
US-08-474-410-3

Alignment Scores:
Pred. No.: 3.06e-50 Length: 1255
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 3 Gaps: 10

US-09-208-629f-6 (1-408) x US-08-474-410-3 (1-1255)
Qy 80 GlyAlaThrIleThrValLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db 212 GGAGTTACAGTTGAACAGCTCTTTCTGTGGATGAGTTTCTGCATCT-----259
Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119

Db 260 -----GTCCCTCACTGGAACACTGACCACCTGTCTCTCCCAATTGTC 301
Qy 120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 302 TACACAAATTGTTGTGGTGGTTTCCCAAGTAAGCGCATGGCCCTGTGGGTCTTCTT 361
Qy 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCCGAACCTAAGAGAAGCACCCTGCTGTGATTATACATGGCCAATCTGGCCCTGGCTGAC 421
Qy 159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db 422 CTCCTCTCTGTCTATCTGGTCCCTTGAAGATTGCTATCATCATCATGCAACCACTGG 481
Qy 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 482 ATTATGGGAAGCTCTTTGTAAATGTCTTATGGCTTTTCTATGGCAACATGACTGT 541
Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 542 TCCATTCTCTCATGACCTGCCTCAGTGTGAGAGTATTGGGTCTATCGTGAACCCCATG 601
Qy 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 602 GGGCAC---TCCAGGAAGAAGCAACATTTGCCATTGGCATCTCCCTGGCAATATGGCTG 658
Qy 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 659 CTGATTCTGCTGTGCACCATCTCTTGTATGTGTGGAAGCAGCACCATTCTTCATCTCTGCC 718
Qy 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln 278
Db 719 CTGAACATCACGACCTGTCATCATGTT-----TTGCCTGAGCAG 757
Qy 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db 758 CTCCTTGGTGGGAGACATGTTCAATTACTTCTCTCTCTGGCCATTGGGGTCTTTCTGTTC 817
Qy 293 ProPheValLeuIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----310
Db 818 CCAGCCTTCTCCACAGCCTCTCCCTATGTGTGATGATCATGAATGCTGGCATCTTCTGCC 877
Qy 311 -----TyrAspHisArgTrpLeuTyrTyrValLysAlaSerLeuLeuLeu 326
Db 878 ATGGATGAAACTCAGAGAAGAAAGGAGAGGCCATCAAACTCATTTCTACTGTCCTG 937
Qy 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn 346
Db 938 GCCATGTACCTGATCTGCTTCACTCTAGTAACCTTCTGTTGTGTGTCAT-----988
Qy 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAla 363
Db 989 TATTTTCTGATTAAGAGCCAGGCGCAGGCCATGCTATGCCCCTGTACATTTAGCCCTC 1048
Qy 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 1049 TGCCTCTTACCCTTAACAGCTGATCGACCCCTTTGTCTATCTACTTTGTTTCACATGAT 1108
Qy 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGATCATGCAAGAAGACGCTCTCCITTCGCGAAGTGTCGCCGACGTGTAAGCAG 1165
RESULT 21
US-08-486-673B-3
; Sequence 3, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B

```

; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1255
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (56)..(1249)
; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein
; OTHER INFORMATION: sequences
US-08-486-673B-3

Alignment Scores:
Pred. No.: 3,06e-50 Length: 1255
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 4 Gaps: 10

US-09-208-629F-6 (1-408) x US-08-486-673B-3 (1-1255)

QY 80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db 212 GGAGTTACAGTTGAAACAGCTTTCTGTGGATGAGTTTCTGCACT-----259
QY 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIle 119
Db 260 -----GTCCCTCACTGGAAACTGACCACCTGCTTCCTCCATTCGCAATGTC 301
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 302 TACACAAATGTGTTGGTGGGTTTGGCCAGTAACGGATGGCCCTGTGGGTCTTCTT 361
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCGGAACTAAGAAGAAGACACCTGCTGTGATTTACATGGCCAATCTGGCCTTGGCTGAC 421
QY 159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnTrp 178
Db 422 CTCCTCTCTGTCATGGTTCCCTTGAAGATTGCCTATCACATGCGCAACAACACTGG 481
QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 482 ATTTATGGGAAGCTCTTTGATATGTGCTTATTGGCTTTTCTATGGCAACATGACTGT 541
QY 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrIleuAlaIleValHisProPhe 218
Db 542 TCCATTCTCTCATGACCTGCTCAGTGTGAGAGTATGGTGCATCGTGAACCCCATG 601
QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db 602 GGGCAC---TCCAGGAAGAGGCAAAATGCGCATCTCCCTGGCAATATGCGTG 658
QY 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 659 CTGATTCGTGTCACCATCCCTTTGTATGCTGAGAGCAGACCATCTTCATCTCTGCC 718
QY 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 719 CTGAACATCAGCACCTGTCATGATGT-----TTCGCTGAGCAG 757
QY 279 LeuTyr-----TyrPheIleSerIleuAlaPhePheGlyPheLeuIle 292
Db 758 CTTCTGGTGGGAGACATGTTCAATTTCTCTCTCTGCGCCATGCGGTCTTTCTGTTTC 817
QY 293 ProPheValIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----310

```

```

Db 818 CCAGCCTTCTCACAGCCTCTGCCTATGTGCTGATGATCAGAACTGTCGATCTTCTGCC 877
QY 311 -----TyrAspHisArgTrpLeuTyrValLysAlaSerLeuLeuLeu 326
Db 878 ATGGATGAAACTCAGAGAAAGAGGAGGCGCATCAAACTCATTTGCTCACTGTCCTG 937
QY 327 ValIlePheThrIleCysPheAlaProSerAsnIleLeuIleIleHisAlaAsn 346
Db 938 GCCATGTACTGATCTGTTCTACTCTAGTAACCTTCTGCTTGTGTGAT-----988
QY 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 989 TATTTTCTGATTAAGAGCGGCGCAGAGCCATGCTATGCCCTGTACATTTAGCCCTC 1048
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 1049 TGCCTCTCTACCCCTTAACAGCTGCATGACCCCTTGTCTATTACTTTGTTTCACATGAT 1108
QY 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGGATCATGAAAGAACGCTCTCTTCCGAGTGTCGCGACTGTAAAGCAG 1165

RESULT 22
US-08-476-000-62
; Sequence 62, Application US/08476000
; Patent No. 5716789
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,000
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..1240
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 50
US-08-476-000-62

```


Alignment Scores:

Pred. No.: 7.59e-49 Length: 1414
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 1 Gaps: 14

US-09-208-629f-6 (1-408) x US-08-472-840-62 (1-1414)

QY 8 Ser***ArgLeuArgaspGlyThrGlnValIleLysMetLysAlaLeuIlePheAla 27
Db 41 TCCAGGAGGATCGGAGCCCGCGGGTGGCTGCTGGGGCGCCCATCTCTGTAGCA 100
QY 28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db 101 GCCTCTCTC-----TCTGTC---AGTGGCACCATTCCCAAGAACCAAT 139
QY 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProAsnSerPhe 67
Db 140 AGATCTCTAAAGGAAGAGCCTTATGTGTAGGTTGATGGACA----- 184
QY 68 GluGluPheProPheSerAlaLeuGluGlyThrGlyAlaThrIleThrValLysIle 87
Db 185 -----TCCACAGTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226
QY 88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db 227 TGTGTGATGAGTTTCTGCATCT-----GTCCTC 256
QY 108 ThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeuLeuValPheValGly 127
Db 257 GCTGGAAACTGACCACTGCTCTCCATTCCTCAATGTCTACAAATGTTGCGGTGGT 316
QY 128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db 317 TTGCCAAGTAACGGCATGGCCCTATGGGTCTTTTCCGAACCTAAGGAAGACCCCT 376
QY 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 377 GCTGTGATTTACATGCGCAATCTGGCTTGGCTGACCTCTCTCTCATCTGTTCCCTC 436
QY 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 437 TTCAAGATTGCCTATCACATACATGCGCAACATGGATTATGAGGAAGCTCTTTGTA 496
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 497 GTGCTTATGGCTTTTCTATCGCAACATGATCTGTTCCATTCTTTCATGACCTGCCTC 556
QY 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db 557 AGTGTGCGAGGATTTGGGTGTCATCGTGAACCCCATGGGCAC---TCCAGGAAGAGCA 613
QY 227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db 614 AACATTGCCATGGCATCTCCCTGGCAATATGGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
QY 247 PhePheIleLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db 674 TTGTATGTCGTGAAGCAGACCATCTTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
QY 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr----- 280
Db 734 GTT-----TTGCCTGACAGCTCTTGTGGGAGACATGTTCAAT 772
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuProPheValLeuIleIleTyrCys 300
Db 773 TACTTCTCTCTGCGCCATGGGGTCTTCTGTTCCAGCCTTCTCTCAGCCTCTGCC 832
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 833 TATGTGCTGATGATCAGATGCTGCGATCTTCTGCTGATGATGATGATGATGATGATGAT 892

Alignment Scores:

QY 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db 893 AGGAAGAGGGCCATCAAACTATTGTCTACTGTCTGGGCATGCTACTGATCTGCTTCACT 952
QY 335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 953 CCTAGTAACCTTCTGCTTGTGTGCAT-----TATTTCTGATTAAAGCCAGGCGC 1003
QY 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1004 CAGAGCCATGCTATGCCCTGTACATTTAGCCCTCTGCTCTTACCTTTAACAGCTGC 1063
QY 372 LeuAspPropheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db 1064 ATCGACCCCTTGTCTATTACTTTGTTTCATGATTTTCAGGATCATGCAAGAACGCT 1123
QY 390 TyrLeuThrLys*****AsnAspLeuArgGluGlnGlnPro 404
Db 1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168

RESULT 24
US-08-476-976-62
; Sequence 62, Application US/08476976
; Patent No. 5874400
; GENERAL INFORMATION:
; APPLICANT: SUNDELIN, JOHAN
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/476,976
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,301
; FILING DATE: 25-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ADLER, REID G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 2803-0006.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 62:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1414 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 50..1240
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 50
; US-08-476-976-62

Pred. No.: 7.59e-49 Length: 1414
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 14 Gaps: 14

US-09-208-629f-6 (1-408) x US-08-476-976-62 (1-1414)

QY 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db 41 TCCAGGAGGATCGGAGCCCGCGGCTGCTGGGGCCCGCATCTCTGTAGCA 100
QY 28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db 101 GCCTCTCTC-----TCCCTGC---AGTGGCACCATTCCAAAGAACCAAT 139
QY 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db 140 AGATCTCTTAAGAGAGAGAGCTTATTTGGTAAGTTGATGGGCACA-----184
QY 68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
Db 185 -----TCCACGTCACCTGGA---AAAGGAGTTACATTTGAACAGTCTTT 226
QY 88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTrpLeu 107
Db 227 TCTGTGCATCAGTTTCTGCATCT-----GTCCCTC 256
QY 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
Db 257 GCTGGAAGAGTACCACTGTCTTCCATTTGCTACACAAATGTGTTCGGTGGGT 316
QY 128 ValProAlaAsnAlaValThrLeuTrpMetLeuPheArgThrArgSerIleCysThr 147
Db 317 TTGCAAGTAACGGCATGCCCTATGGCTTCTTTTCGAACTAAGAGAGCACCT 376
QY 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 377 GCTGTGATTTACATGCCAATCTGGCTTGGCTGACCTCTCTCTGTCTGTTTCCCC 436
QY 167 PheLeuAlaIleAlaThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 437 TTGAAGATTGCCCTATCACATACATGCAACAACTGGATTATGGGAAGCTCTTTGTAAT 496
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 497 GTGCTTATTTGGCTTTTCTATCGCAACATGACTGTTCATTTCTTCATGACCTGCCCTC 556
QY 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db 557 AGTGTGCAGAGGTATTGGGTCTATCGTGAACCCCATGGGCAC---TCCAGGAAGAGGCA 613
QY 227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db 614 AACATTGCCATTGGCATCTCCCTGGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 673
QY 247 PhePheIleLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db 674 TTGTATGTCGTAGAGCAGACCATCTTCATCTCGCCCTGAAACATCACGACCTGTCTATGAT 733
QY 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr-----280
Db 734 GTT-----TTGCTGTGACGCTCTTGTGGGAGACATGTTCAAT 772
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValleuIleTyrCys 300
Db 773 TACTTCTCTCTGGCCATCGGGTCTTTCTTCTTCCAGCCTTCTCTACAGCCTCTGCC 832
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 833 TATGTCTGATCATCAGATGCTGCGATCTTGTGCCATGGATGAACAACTCAGAGAGAAA 892

QY 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db 893 AGGAAGAGGGCCATCAAACTATGTCACTGCTCTGGGCATGTACCTGATCTGCTTCACT 952
QY 335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrAsnAsnThrAspGly 354
Db 953 CCTAGTAACCTTCTGCTTGTGTGCAT-----TATTTTCTGATTAAAGACCCAGGCG 1003
QY 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1004 CAGAGCCATGTATGCCCTGTACATGTAGCCCTCTGCTCTTACCTTTAACAGCTGC 1063
QY 372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db 1064 ATCGACCCCTTTGTCTATTACTTTTTCACATGATTTTCAGGATCATCGAAAGACGCT 1123
QY 390 TyrLeuThrLys*****AsnAspLeuArgGluGlnGlyGlnPro 404
Db 1124 CTCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1168
RESULT 25
US-08-474-410-62
Sequence 62, Application US/08474410
Patent No. 6043212
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,410
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 1414 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 50..1240
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 50
US-08-474-410-62
Alignment Scores: 7.59e-49 Length: 1414
Pred. No.:

```
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 3 Gaps: 14

US-09-208-629F-6 (1-408) x US-08-474-410-62 (1-1414)

QY 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaA 27
DB 41 TCCAGAGGATCGGAGCCCGCGCGGTGGCTGCTGGGGCCGCCATCTCTGTAGCA 100
QY 28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
DB 101 GCCTCTCTC-----TCTGTC---AGTGGCACCATCCAGGAACCAAT 139
QY 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
DB 140 AGATCTCTAAAGGAAGAGCCCTTATTGTAAGTTGATGGCACA-----184
QY 68 GluGluPheProPheSerAlaLeuGluGlyTTPThrGlyAlaThrIleThrValLysIle 87
DB 185 -----TCCACGCTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226
QY 88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
DB 227 TCTGTGATGAGTTTCTGCATCT-----GTCCTC 256
QY 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
DB 257 GCTGGAAACCTGACCACTGTCTTCCATTTGCTACCAATTTGTTGGGTGGGT 316
QY 128 ValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIleCysThr 147
DB 317 TTCCCAAGTAACGGCATGGCCCTTGTCTTTTCCGAACCTAAGGAAGACACCT 376
QY 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
DB 377 GCTGTGATTACATGGCCATCTGGCTTGGCTGACTCTCTCTCATCTGTTCCCC 436
QY 167 PheIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValLeuCysArg 186
DB 437 TTGAAGATTGCCATCACATACATGCAACCACTGGATTATGGAAGCTCTTTGTAAT 496
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
DB 497 GTCCTATTGGCTTTTCTTCCCAACATGACTGTCTTCCATCTCTTCAGACCTGCC 556
QY 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
DB 557 AGTGTGCAGAGGTATTGGGTCTATCGTGAACCCCATGGGCAC---TCCAGGAAGAGCA 613
QY 227 ThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeuTyrMetLeuPro 246
DB 614 AACATTGCCATTTGGCATCTCCCTGGCAATATGGCTGCTGACTGTGTGTGTCACCATCT 673
QY 247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
DB 674 TTGTATGTCTGTAAGACAGACCATCTTCATTTCTGCGCCCTGAACATCAGACCTGTCA 733
QY 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr-----280
DB 734 GTT-----TTGCTGACAGCTCTTGTGGGGAGACATGTTCAAT 772
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyrCys 300
DB 773 TACTTCTCTCTGTGGCCATTGGGGTCTTCTTCTCCAGCCCTTCTCAGCCCTCTGCC 832
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
DB 833 TATGTCGTGATGATGAGATGCTGCGATCTTCTGCCATGATGAAACCTCAGAGAGAA 892
QY 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
```

RESULT 26

```
US-08-486-673B-62
; Sequence 62, Application US/08486673B
; Patent No. 6297026
; GENERAL INFORMATION:
; APPLICANT: Scarborough, Robert M.
; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor
; FILE REFERENCE: 44481-5006-08-US
; CURRENT APPLICATION NUMBER: US/08/486,673B
; CURRENT FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/097,938
; PRIOR FILING DATE: 1993-07-26
; PRIOR APPLICATION NUMBER: PCT/US94/08536
; PRIOR FILING DATE: 1994-07-26
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 62
; LENGTH: 1414
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (50)..(1240)
; OTHER INFORMATION: C140 receptor, cDNA and deduced protein sequences
US-08-486-673B-62
```

```
Alignment Scores:
Pred. No.: 7,59e-49 Length: 1414
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 4 Gaps: 14
```

US-09-208-629F-6 (1-408) x US-08-486-673B-62 (1-1414)

```
QY 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaA 27
DB 41 TCCAGAGGATCGGAGCCCGCGCGGTGGCTGCTGGGGCCGCCATCTCTGTAGCA 100
QY 28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
DB 101 GCCTCTCTC-----TCTGTC---AGTGGCACCATCCAGGAACCAAT 139
QY 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
DB 140 AGATCTCTAAAGGAAGAGCCCTTATTGTAAGTTGATGGCACA-----184
QY 68 GluGluPheProPheSerAlaLeuGluGlyTTPThrGlyAlaThrIleThrValLysIle 87
DB 185 -----TCCACGCTCACTGGA---AAAGGAGTTACAGTTGAAACAGTCTTT 226
QY 88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
```



```
Db 1212 AACGTCCTCTGATTGCGGATTACTACTTCTTCTCACACTTCCACACAGAGGCTGCC 1271
QY 356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
Db 1272 TACTTTGGCTACTCTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
QY 376 LeuTyrPheLeuMetSer 381
Db 1332 ATTACTATTACGCTTCC 1349

RESULT 31
US-08-475-263-219
; Sequence 219, Application US/08475263
; Patent No. 575994
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: MORRISON & FORSTER
; STREET: 2000 Pennsylvania Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,263
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
US-08-475-263-219

Alignment Scores:
Pred. No.: 2 96e-43 Length: 3480
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: 1 Gaps: 11

US-09-208-629f-6 (1-408) x US-08-475-263-219 (1-3480)
QY 14 GlyThrGlnValIleLeuMetCysLeuAlaLeuPheAlaLeuLeuLeuLeuLeu 33
Db 228 GGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db 288 GCCCGCACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
```

```
QY 54 LeuProIleLeuLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db 336 TTATGATCCCGGTCTCTTCTCTCAGGAACCCCAATGATAAATGAACCATTTTGGGAG 395
QY 73 SerAlaLeuGluGlyTyrThrGlyAlaThr-----IleThrValIleLys 88
Db 396 GATGAGGAGAAAATGAAGTGGTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
QY 89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
Db 456 AGTCCTCTTCAAAACAACCTTCTCCATTCATCTCAGAGATGCTCCGGATATTGACC 515
QY 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGlyVal 128
Db 516 AGCTCCTGGCTGACACTCTTTGTCCCATCTGTGTACACGGAGTGTGTGTGTGTGTGTGTGT 575
QY 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
Db 576 CCACCTAAACATCATGGCCATCGTGTGTTCATCTGAAATGAAGTCAAGAAG---CG 632
QY 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 633 GCGGTGTGTACATGCTGCACCTGCGCAGGAGATGTCTGTGTGTGTGTGTGTGTGTGTGTGT 692
QY 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGGTGTAATTGTGTGCG 752
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
Db 753 TTCGTCACTGCAGCAATTTTACTGTAAACATGTACGCTCTATCTTCTCATGACAGTCATA 812
QY 207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
Db 813 AGCATGACCGGTTTCTGGCTGTGTGTGTATCCATGCAGTCCCTCTCTGGGCTACTCTG 872
QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
Db 873 GGAAGGCTTCCTTC-----ACTTGTCTGGCCATCTGGGCTTTGGCCATCGCAGG 923
QY 244 MetLeuProPhePheIleLeuLysGlnTyrTyrLeuValGlnProAspIleThrThr 263
Db 924 GTAGTGCCTCTCGTCTCAAGGAGCAAAACCATCCAGGTGCCCGGCTCAACATCACTACC 983
QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
Db 984 TGTCAATGATGCTCAATGAAACCTGCTCGAAGC---TACTATGCTTACTACTTCTCA 1040
QY 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaLa 303
Db 1041 GCCTTCTCTGTGTCTCTCTTTTGTGGCGGTGATCATTTCCACGGTCTGTATGTGTCT 1100
QY 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrpLeu 316
Db 1101 ATCATTCGATGTCTTAGTCTTCCAGTTCGCAACCGCAGCAAGAAGTCCCGGCTTTG 1160
QY 317 TrpTyrValLysAlaSerLeuLeuValIlePheThrIleCysPheAlaProSer 336
Db 1161 TTC-----CTGTCACTGCTGTCTTTTCTGCATCTTTCATTTCTTCGACCCACA 1211
QY 337 AsnIleIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
Db 1212 AACGTCCTCTGATTGCGCATTTACTTCTTCTTCTCACACTTCCACACAGAGGCTGCC 1271
QY 356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
Db 1272 TACTTTGGCTACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
QY 376 LeuTyrPheLeuMetSer 381
Db 1332 ATTACTATTACGCTTCC 1349
```

RESULT 32

US-08-485-886-219
; Sequence 219, Application US/08485886
; Patent No. 5798248
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-485-886-219

Alignment Scores:
Pred. No.: 2,96e-43 Length: 3480
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: 1 Gaps: 11

- US-09-208-629F-6 (1-408) x US-08-485-886-219 (1-3480)

Qy 14 GlyThrGlnValIleLeuMetLysAlaLeuPheAlaAlaGlyLeuLeuLeu 33
Db 228 GGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 287
Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThr 53
Db 288 GCGCGCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 335
Qy 54 LeuProIleLysThrPheArgGlyAlaProAsn---SerPheGluGluPheProPhe 72
Db 336 TTATGATCCCGGCTATTTCTTCTCAGGAACCCCAATGATAATATGACCATTTGGGAG 395
Qy 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValLysLys 88
Db 396 GATGAGGAGAAAGTAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455

Qy 89 CysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
Db 456 AGTCCTCTTCAAAACAACTTCTCCATCATCTCAGAGATGCGCTCGGATATTGACC 515
Qy 109 SerSerLeuSerThrLysLeuLeuProAlaIleTyrLeuValPheValValGlyVal 128
Db 516 AGCTCTCGGCTGACACTCTTTGTCCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCCTC 575
Qy 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
Db 576 CCACATAAACATCATGCGCATCGTTGTGTCTCATCTCCGAAATGAAGGTCAAGAAG--CG 632
Qy 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 633 GCGGTGGTGTACATGCTGACCTGCGCAGGAGATGTCTGTGTGTGTGTGTGTGTGTGTGT 692
Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 693 TTTAAGATCAGCTATTACTTTTCCGCGAGTGTATGCGAGTTTGGGTCTGAATTGTGTGCG 752
Qy 187 AlaThrThrValIlePheTyrClyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 753 TTGTCACCTGCAGCATTTTACTGTAAACATGTACGCTCTATCTGTCTCATGACAGTCATA 812
Qy 207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
Db 813 AGCATGTACCGTTTCTGCTGTGTGTGTATCCCATGCGAGTCCCTCTCTGCGGCTACTGT 872
Qy 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
Db 873 GGAAGGGCTTCCTTC-----ACTGTCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923
Qy 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
Db 924 GTAGTGCCTCTGCTCTCTCTTTTGTGCGGCTGATCAATTCACGGCTGTGTATGTGTCT 983
Qy 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
Db 984 TGTCTATGATGCTCTCAATGAACCTGCTCGAAGG---TACTATGCTCTACTACTCTCTCA 1040
Qy 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
Db 1041 GCCTTCTCTGCTGCTCTCTTTTGTGCGGCTGATCAATTCACGGCTGTGTATGTGTCT 1100
Qy 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrpLeu 316
Db 1101 ATCATTCGATGCTTATGCTCTTCCGAGTTCGCAACCGCGAGCAAGGTCCCGGGCTTTG 1160
Qy 317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
Db 1161 TTC-----CTGTCAGCTGCTGTTTCTGCACTTCTCATCTATTGCTTCGACCCACA 1211
Qy 337 AsnIleIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
Db 1212 AACGTCCTCCTGATTGGCATTACTATCTCTTCTCTACACTTCCACACAGAGGCTGCGC 1271
Qy 356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe 375
Db 1272 TACTTTGCTTACTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
Qy 376 LeuTyrPheLeuMetSer 381
Db 1332 ATTACTATTACGCTTCC 1349

RESULT 33

US-08-477-362-219
; Sequence 219, Application US/08477362
; Patent No. 5849507
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND

```

; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,362
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-477-362-219

Alignment Scores:
Pred. No.: 2,96e-43 Length: 3480
Score: 523.00 Matches: 127
Percent Similarity: 76 Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: 2 Gaps: 11

US-09-208-629f-6 (1-408) x US-08-477-362-219 (1-3480)

Qy 14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaGlyLeuLeuLeu 33
Db 228 GGCGCGGGGGCTGCTGCTGGTGGCGGCTGCTGCTGCTGCTGCTGCTGCT 287
Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db 288 GCCCGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335
Qy 54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db 336 TTAGATCCCGGCTATTTCTTCTCAGGAACCCCAATGATAAATATGAAACATTTGGGAG 395
Qy 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValLysIleLys 88
Db 396 GATGAGAGAAAATGAAGTGGTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
Qy 89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
Db 456 AGTCCCTCTCAAAACAACTCTCTGATTCATCTCAGAAAGATCGCTCCGATATTGACC 515
Qy 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGlyVal 128
Db 516 AGCTCTGGTGACACTTTTGTGCTACACCGGAGTGTGTGTAGTACGCCTC 575

; 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 576 CCACATAACATCATGCCCATCGTTGTTGTTTCATCTGAAATGAAGGTCAAGAAG---CCG 632
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 633 GCGGTGGTGTATCATGCTGCACCTGCCACGCGAGATGTGCTGTTGTGTGTGTGCTGCC 692
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 693 TTATAGATCAGCTATTTACTTTCCGCGCAGTATGGCAGTTTGGGTCTGAATTTGTCGC 752
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 753 TTGCTCACTGCAGCATTTTACTGTAAATGTACGCTCTATCTTGTCTCATGACAGCATCA 812
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 813 AGCATGTACCGGTTTCTGGCTGTGTGTATCCCATGCAGTCCCTCTCTGCGGTACTCTG 872
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 873 GGAAGGGCTTCCTC-----ACTGTCTGGCCATCTGGCCTTGGCCATCGCAGGG 923
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 924 GTAGTGCCTCTCGTCTCAAGGAGCAAAACATCCAGGTGCCCGGCTCAACATCACTACC 983
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 264 CysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTyrPheIle 283
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 984 TGTCTATGATGTCTCAATGAAACCCCTGCTCGAAGGC---TACTATGCCCTACTACTTCTCA 1040
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyrCysTyrAlaAla 303
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1041 GCCTCTCTGCTGCTCTTCTTTTGTGCGCTCATCTTCACGGTCTGTATGTGTCT 1100
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTyrLeu 316
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1101 ATCATTCGATGCTTAGCTCTCCGAGTTGCCAGCGCAGCAAGAAGTCCCGGCTTIG 1160
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 317 TrpTyrValLysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSer 336
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1161 TTC-----CTGTCACTGCTGTTTCTGTCATCTTCTCATCTTGTGCGACCCACA 1211
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 337 AsnIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1212 AACGCTCTCTGATTCGCGCATTTCTTCCTTCTCAGCTTCCACACGAGGCTGCC 1271
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 356 TyrPheIleTyrLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe 375
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1272 TACTTTGCCTACCTCTCTGCTGTGTGTGTCAGCAGCATAAGCTCGTGCATCGACCCCTA 1331
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 376 LeuTyrPheLeuMetSer 381
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; 1332 ATTACTATTACGCTTCC 1349
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 34
US-08-477-134-219
; Sequence 219, Application US/08477134
; Patent No. 5856448
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
```

```

; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,134
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 220000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-477-134-219

Alignment Scores:
Pred. No.: 2,96e-43 Length: 3480
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: Gaps: 11

US-09-208-629F-6 (1-408) x US-08-477-134-219 (1-3480)
Qy 14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaAlaGlyLeuLeuLeu 33
Db 228 GGGCGCGGGCTGCTGCTGGTGGCGCTGCTTCACTGTCGCGCGCTGTTGCT 287
Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db 288 GCCCGACCGCGCGCGCGCGCGAGATCAAAAGCAACAAAT-----GCCACC 335
Qy 54 LeuProIleLysThrPheArgGlyAlaProAsn---SerPheGluGluPheProPhe 72
Db 336 TTGATATCCCGGCTATTTCTTCTCAGAACCCCAATGATAAATATGAACCATTTGGGAG 395
Qy 73 SerAlaLeuGluGlyTrpThrGlyValAlaThr-----IleThrValLysIleLys 88
Db 396 GATCAGAGAAAATGAAGTGGTTAACTACATACAGATTAGTCTCCATCAATAAAGC 455
Qy 89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyLeuThr 108
Db 456 AGTCCTCTTCAAAACAACTCTCTGCTATTCATCTCAGAGATGCTCCGAGATATTGACC 515
Qy 109 SerSerLeuSerThrLysLeuPheAlaIleTyLeuLeuValPheValGlyVal 128
Db 516 AGTCCTCTGAGACACTTTTGTCCCATCTGTGTACACCGAGGTGTGTAGTCAGCCTC 575
Qy 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPheArgThrArgSerIleCys 146
Db 576 CCACATAACATCATGGCCCATGTTGTGTCTCATCTCGAAATGAAGGTCAAGAAG---CCG 632
Qy 147 ThrThrValPheTyThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 633 GCGGTGTGTACATGCTGCACCTGGCCACGAGATGCTGTTGTGTGCTGCTGCTCCCC 692
```

```

Qy 167 PheLysIleAlaTyThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 693 TTTAAGATCAGCTATTACTTTTCCGGCAGTGATTGGCAGTTTGGCTGAATTGTGTCG 752
Qy 187 AlaThrThrValIlePheTyThrGlyAsnMetTyThrCysSerIleLeuLeuLeuAlaCysIle 206
Db 753 TTCTGTCACCTGCAGCATTTTACTGTAAACATGTACGCTCTATCTTCTCATGACAGTCATA 812
Qy 207 SerIleAsnArgTyThrLeuAlaIleValHisPro-----PheThrTyArgGlyLeu 223
Db 813 AGCATGTACCGGTTTCTGCTGTGGTGATCCCATGCAGTCCTCTCTCGGGGTACTCTG 872
Qy 224 ProLysHisThrTyThrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTy 243
Db 873 GGAAGGGCTTCCTTC-----ACTTGTCTGGCCATCTGGGCTTGGCCCATCGCAGG 923
Qy 244 MetLeuProPhePheIleLeuLysGlnGlyTyThrLeuValGlnProAspIleThrThr 263
Db 924 GTAGTGCCTCTGCTCTCAAGGAGCAACCATCCAGGTGCCCGGGCTCAACATCACTACC 983
Qy 264 CysHisAspValHisAsnThrCysGluSerSerPropheGlnLeuTyThrPheIle 283
Db 984 TGTATGATGTGCTCAATGAACCTGCTCGAAGG---TACTATGCCCTACTACTTCTCA 1040
Qy 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyThrCysTyAlaAla 303
Db 1041 GCCTTCTGCTGCTCTTCTTTTGTGCGGCTGATCAATTCACCGCTGTTATGTGCT 1100
Qy 304 IleIleArgThrLeuAsn-----AlaTyAspHisArgTrpLeu 316
Db 1101 ATCATTCGATGTCTTAGCTCTTCCGAGTTGCCAAGAGTCCCGGGCTTTG 1160
Qy 317 TrpTyThrValLysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSer 336
Db 1161 TTC-----CTGTACGCTGCTGTTTCTGCACTTCTATCATTTGTTTCGACCCACA 1211
Qy 337 AsnIleIleLeuIleHisAlaAsnTy---TyThrAsnAsnThrAspGlyLeu 355
Db 1212 AAGTCCTCTCATTTGGCATTACTTCTTCTCACACTTCCACACAGAGGCTGCC 1271
Qy 356 TyThrPheIleTyLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProphe 375
Db 1272 TACTTTGCTACCTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
Qy 376 LeuTyThrPheLeuMetSer 381
Db 1332 ATTACTATTACGCTTCC 1349
```

RESULT 35

```

US-08-473-489A-219
; Sequence 219, Application US/08473489A
; Patent No. 6024936
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,489A
```

```

; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/789,184
; FILING DATE: 1991-11-07
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 22000-20502.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 34-0154
; INFORMATION FOR SEQ ID NO: 219:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..1499
; US-08-473-489A-219

Alignment Scores:
Pred. No.: 2,96e-43 Length: 3480
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: 11 Gaps: 11

US-09-208-629f-6 (1-408) x US-08-473-489A-219 (1-3480)

QY 14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaGlyLeuLeuLeu 33
DB 228 GGGCCGGCGGCTGTGTGGCGGCTGTTTCTGAGTGTGGCGGCGGCTGTGTCT 287
QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
DB 288 GCCCGCAGCCGGCGCCGAGGAGCAGAAATCAAAAGCAAAAT-----GCCACC 335
QY 54 LeuProIleLysThrPheArgGlyAlaProAsn---SerPheGluGluPheProPhe 72
DB 336 TTAGATCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAAACATTTTGGGAG 395
QY 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValLysIleLys 88
DB 396 GATGAGAGAGAAATGAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAAGC 455
QY 89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
DB 456 AGTCCTCTTCAAAACAACTTCTGCTATTCATCTCAGAGATGCTCCGATATTGACC 515
QY 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGlyVal 128
DB 516 AGCTCTGGGTGACACTTTTGTGCCATCTGTGTACACCGGAGTGTGTGTAGTCAGCCTC 575
QY 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPheArgThrArgSerIleCys 146
DB 576 CCACATAACATCATGCGCCATCTGTGTGTCACTGAAATGAGGTCAAGAG-----CCG 632
QY 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
DB 633 GCGGTGGGTACATGCTGTCACCTGGCCAGGAGATGTGTGTGTGTGTGTGTGTGTGTGT 692
QY 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
DB 693 TTAAAGATCAGCTATTACTTTTCCCGCAGTATGTCAGTGTGGGTCTGAATTTGTGTGTC 752
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
DB 753 TTCGTCACTGCAGCATTTTACTGTAAACATGTAGCCCTCTATCTTGTCTCATGACGTCA 812

207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
813 AGCAITGACCGGTTTCTGGCTGTGGTGTATCCATGCAGTCCCTCTCTCTGGGCTACTCTG 872
QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
DB 873 GGAAGGGCTTCTTC-----ACTTGTCTGGCCATCTGGGCTTTGGCCATTCGAGGG 923
QY 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
DB 924 GTAGTGCCTCTCGTCTCAAGAGCAACCATCCAGGTGCCGGGGCTCAACATCACTACC 983
QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
DB 984 TGTCTATGATGTCTCAATGAAACCCCTGCTCGAAGG---TACTATGCTTACTTCTCTCA 1040
QY 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
DB 1041 GCCTTCTCTGCTGTCTCTTTTGTGGCGGTGATCATTTCCACGGTCTGTGTATGTGTCT 1100
QY 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrpLeu 316
DB 1101 ATCATTCGATGTCTTAGCTTCCGAGTTGCCAAGGAGTCCCGGGCTTTG 1160
QY 317 TrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
DB 1161 TTC-----CTGTCAGTCTGTGTTTCTGCACTTTCATCATTTGCTTCGAGCCACCA 1211
QY 337 AsnIleIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
DB 1212 AACGTCCTCTGATTTGCGGCTTACTTCTTCTCAGCTTCCACCATCCACACAGAGGCTGCC 1271
QY 356 TyrPheIleTyrLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
DB 1272 TACTTTGCTACTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1331
QY 376 LeuTyrPheLeuMetSer 381
DB 1332 ATTACTATTACGCTTCC 1349

RESULT 36
US-08-485-695-219
; Sequence 219, Application US/08485695
; Patent No. 6124101
; GENERAL INFORMATION:
; APPLICANT: COUGHLIN, SHAUN R.
; APPLICANT: SCARBOROUGH, ROBERT M.
; TITLE OF INVENTION: RECOMBINANT THROMBIN RECEPTOR AND
; TITLE OF INVENTION: RELATED PHARMACEUTICALS
; NUMBER OF SEQUENCES: 223
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,695
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/789,184
; FILING DATE: 07-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
```


; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 225..1499
 ; US-08-018-760-219

Alignment Scores:

Pred. No.: 2,96e-43 Length: 3480
 Score: 523.00 Matches: 127
 Percent Similarity: 52.59% Conservative: 76
 Best Local Similarity: 32.90% Mismatches: 153
 Query Match: 24.25% Indels: 30
 DB: 4 Gaps: 11

US-09-208-629f-6 (1-408) x US-08-018-760-219 (1-3480)

QY 14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaLaglyLeuLeuLeu 33
 Db 228 GGGCCCGCGGGCTGCTGCTGGCGCGCTGCTTCCAGTCTGTGCGCGCGCTGTTGTCT 287
 QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
 Db 288 GCCCGCACCGCGCCCGCGCGCGAGCCAGCAATCAAAAGCAACAAAT-----GCCACC 335
 QY 54 LeuProIleLysThrPheArgGlyAlaProAsn---SerPheGluGluPheProPhe 72
 Db 336 TTATATCCCGGTCATTTCTTCCAGAACCCCAATGATAAATATGACCAATTTGGGAG 395
 QY 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrValIleLys 88
 Db 396 GATGAGGAGAAATAAGAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
 QY 89 CysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
 Db 456 AGTCCTCTTCAAAAACAACTCTCGCATTCATCAGAGATGCCCTCGGATATTGACC 515
 QY 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheValValGlyVal 128
 Db 516 AGCTCTGGGTGACACTTTGTGCCCATCTGTGTACCGGAGTGTGTGTAGTCACGCTC 575
 QY 129 ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
 Db 576 CCACATAACATCATGCGCATCTGTGTCTGTTTCACTGAAATGAAGTCAAGAAG---CCG 632
 QY 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
 Db 633 CGCGTGGTGTACATGTCGACCTGGCCGCGCAGATGTGCTGTTGTGTGCTGCTCCCG 692
 QY 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
 Db 693 TTATAGATCAGTATTACTTTCCGCGAGTGTGCGAGTTGGGTCTGAATGTGTGCG 752
 QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIle 206
 Db 753 TTCTGCTCACTGCAGCATTTTACTGTAACATGTACGCTCTATCTGTCTCATGACAGCATA 812
 QY 207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
 Db 813 AGCATTTGACCGGTTTCTGGTGTGGTGTATCCCATGCGATGCTCTCTCTGGCGTACTCTG 872
 QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
 Db 873 GGAAGGCTTCTTCT-----ACTGTGCTGGCCATCTGGGCTTTGGCCATCGCAGGG 923
 QY 244 MetLeuProPhePheIleLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
 Db 924 GTAGTGCCTCTGCTCAAGGAGCAACCATCCAGGTGCGCGGCTCAACATCACTACC 983
 QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
 Db 984 TGTGATGATGTGCTCAATGAAACCCCTGCTCGAAGGC---TACTATGCTACTACTTCTCA 1040

QY 284 SerLeuAlaPhePheGlyPheLeuIlePheProPheValLeuIleIleTyrCysTyrAlaAla 303
 Db 1041 GCCTTCTCTGCTGCTCTTCTTTTGTGCGGCTGATCATTTCCACGGCTGTTATGTGTCT 1100
 QY 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrpLeu 316
 Db 1101 ATCATTCGATGCTTAGCTCTTCCAGTGTCCACCGCAGCAAGAAGTCCCGGGCTTTG 1160
 QY 317 TrpTyrValLysAlaSerLeuLeuIlePheThrIleCysPheAlaProSer 336
 Db 1161 TTC-----CTGTCAGCTGCTGTTTCTGCACTTTCATCATTTGCTTCGACCCACA 1211
 QY 337 AsnIleIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
 Db 1212 AAGTCTCTCTGATTCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1271
 QY 356 TyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
 Db 1272 TACTTGTGCTACT 1331
 QY 376 LeuTyrPheLeuMetSer 381
 Db 1332 ATTACTATTACGCTTCC 1349
 RESULT 38
 US-08-313-553-12
 ; Sequence 12, Application US/08313553
 ; Patent No. 5641650
 ; GENERAL INFORMATION:
 ; APPLICANT: TURNER, George J.
 ; APPLICANT: BETLACH, Mary C.
 ; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
 ; IN HALOBACTERIA
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Walter H. Dreger
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/313,553
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/038,662
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dreger, Walter H.
 ; REGISTRATION NUMBER: 24,190
 ; REFERENCE/DOCKET NUMBER: A-57669/WHI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1764 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: repeat_region
 ; LOCATION: 436..462
 ; OTHER INFORMATION: /note= "Sequence encoding
 ; Patent No. 5641650

Db 1372 TTCTTTCACACTCCACAGAGGCGCTACTTTGGCCCTACCTCTCTGTGTCTGT 1431
QY 365 LeuGlySerLeuAsnSerCysLeuAspPropheLeuTyPheLeuMetSer 381
Db 1432 GTCAGCAGGATAGCTCGTGCATCGACCCCTAATTACTATTACGCTTCC 1482

RESULT 39
US-08-767-993-12
Sequence 12, Application US/08767993
Patent No. 6010885
GENERAL INFORMATION:
APPLICANT: TURNER, George J.
APPLICANT: BETLACH, Mary C.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS POLYPEPTIDES
TITLE OF INVENTION: IN HALOBACTERIA
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Walter H. Dreger
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/767,993
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-57669/WH
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1764 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: repeat_region
LOCATION: 436..462
OTHER INFORMATION: /note= "Sequence encoding
Patent No. 6010885
OTHER INFORMATION: polyaspartic acid."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 463..465
OTHER INFORMATION: /note= "Codon encoding the
OTHER INFORMATION: N-terminal amino acid of the human thrombin
OTHER INFORMATION: receptor protein."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1630..1632
OTHER INFORMATION: /note= "Codon encoding the
OTHER INFORMATION: C-terminal amino acid of the human thrombin
OTHER INFORMATION: receptor protein."
FEATURE:
NAME/KEY: repeat_region
LOCATION: 1633..1650
OTHER INFORMATION: /note= "Sequence encoding
Patent No. 6010885
OTHER INFORMATION: polyhistidine."
FEATURE:
NAME/KEY: misc_feature

LOCATION: 648..656
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 893..898
OTHER INFORMATION: /note= "Deleted PstI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1301..1309
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1394..1402
OTHER INFORMATION: /note= "Deleted AlwNI restriction
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: misc_signal
LOCATION: 374
OTHER INFORMATION: /note= "RNA start site."
FEATURE:
NAME/KEY: mutation
LOCATION: replace(1671, "")
OTHER INFORMATION: /note= "C to T mutation removes
OTHER INFORMATION: AlwNI site."
FEATURE:
NAME/KEY: CDS
LOCATION: 376..1650
FEATURE:
NAME/KEY: misc_feature
LOCATION: 376..414
OTHER INFORMATION: /note= "Bacteriorhodopsin
OTHER INFORMATION: pre-sequence."
FEATURE:
NAME/KEY: terminator
LOCATION: 1672..1674
OTHER INFORMATION: /note= "Bacteriorhodopsin stop
OTHER INFORMATION: codon."
FEATURE:
NAME/KEY: terminator
LOCATION: 1651..1653
OTHER INFORMATION: /note= "Thrombin stop codon."
US-08-767-993-12

Alignment Scores:
Pred. No.: 5,05e-43 Length: 1764
Score: 516.50 Matches: 123
Percent Similarity: 53.05% Conservative: 77
Best Local Similarity: 32.63% Mismatches: 144
Query Match: 23.95% Indels: 33
DB: 3 Gaps: 11

US-09-208-629F-6 (1-408) x US-08-767-993-12 (1-1764)

QY 30 LeuLeuLeuLeuLeuProPheCysGln-----SerGlyMet 42
Db 376 ATGTTGGAGTTATTCACACAGCAGTGGAGGGGTATCGCAGGCCAGATCCAGGCGCTG 435
QY 43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAla 62
Db 436 GACTACAAGGAGCATGATGACGTCAGCGCCACTTAGATCCCGGGTCATTCTCTCAGG 495
QY 63 ProProAsn---SerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
Db 496 AACCCCAATGATAAATATGACCATTTTGGGAGGATGAGGAGAAAAAATGAAAGTGGGTTA 555
QY 82 Thr-----IleThrValLysIleLysCysProGluGluSerAlaSerHisLeu 97
Db 556 ACTGAATACAGATTAGTCTCCATCAATAAAACAGTCCTCTTCAAAAACAACCTCTCGCA 615
QY 98 HisValLysAsnAlaThrMetGlyTyLeuThrSerSerLeuSerThrLysLeuIlePro 117

```
Db 616 TTCTATCTCAGAGATGCTCCCGGATATTTGACGAGCTCTGGTGTGACACTCTTTGTGCCA 675
Qy 118 AlaIleTyrLeuLeuValPheValGlyValProAlaAsn-----AlaValThrLeu 135
Db 676 TCTGTGTACACCGAGTGTGTGTAGTCAGCTCCCACTAAACATCATGCGCATGTGTGTG 735
Qy 136 TrpMetLeuPhePheArgThrArgSerIleCysThrValPheTyrThrValPheTyrThr 155
Db 736 TTCTATCTGAAATGAAGGTCAAGAG---CCGGCGGTGTGTACATGTGTCACCTGGCC 792
Qy 156 IleAlaAspPheLeuPheCysValThrLeuProPhePheIleAlaTyrHisLeuAsnGly 175
Db 793 ACGGCAGATGTCTGTTGTGTCTGTCTCCCTTTAAGATCAGCTATTACTTTCCGGC 852
Qy 176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrValPheThrValPheTyrGlyAsn 195
Db 853 AGTGATTGGCAGTTGGGTCTGAATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912
Qy 196 MetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleVal 215
Db 913 ATGTAGCCTCTATCTGCTCATGACAGCTATGAGCATGACCGGTTCTGGCTGGTG 972
Qy 216 HisPro-----PheThrTyrArgGlyLeuProIlePheLeuValThr 232
Db 973 TATCCCATGAGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023
Qy 233 CysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGln 252
Db 1024 TGTCTGCCATCTGGGCTTTGGCCATGCGAGGCTAGTGCCTCTGCTCTCAAGGAGCAA 1083
Qy 253 GluTyrTyrLeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGlu 272
Db 1084 ACATCAGGTGCGCGGCTCAATCAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1143
Qy 273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db 1144 CTCGAGGC---TACTATGCTACTACTCTCTGAGCTTCTGCTGCTGCTGCTGCTGCTG 1200
Qy 293 ProPheValLeuIleTyrCysTyrAlaAlaIleArgThrLeuAsn----- 309
Db 1201 CCGGTGATCTTCCACGGCTGTGTATGTGTGTATGTGTATGTGTATGTGTATGTGTATG 1260
Qy 310 -----AlaTyrAspHisArgTyrLeuTyrTyrValLysAlaSerLeuLeuIle 325
Db 1261 GTTGCCAAACCGCAGCAAGAGTCCGGGCTTTGTTC-----CTGTGAGCTGTGT 1311
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAla 345
Db 1312 TTCTGCATCTTCATCTTCTGCTCGGACCCACAAAGCTCTCTGATGCGCATTA 1371
Qy 346 AsnTyr---TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCys 364
Db 1372 TTCCTTTCTCACATCTCCACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1431
Qy 365 LeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1432 GTACAGCATAGTCTGTCATGTCAGCCCTTAATTTACTATTACGCTTCC 1482
```

RESULT 40

```
US-08-383-750-3
; Sequence 3, Application US/08383750
; Patent No. 5744301
; GENERAL INFORMATION:
; APPLICANT: Birkenbach, Mark
; APPLICANT: Kieff, Elliot
; TITLE OF INVENTION: Epstein Barr Virus Induced Genes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1100 New York Avenue, N.W.,
; STREET: Suite 600
; CITY: Washington
```

```
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,750
FILING DATE: Herewith
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: Fox, Samuel, L.
REGISTRATION NUMBER: 30,353
REFERENCE/DOCKET NUMBER: 0627,3300001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1643 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 34..1116
US-08-383-750-3

Alignment Scores:
Pred. No.: 2,61e-32 Length: 1643
Score: 410.50 Matches: 99
Percent Similarity: 47.8% Conservative: 58
Best Local Similarity: 30.18% Mismatches: 114
Query Match: 19.03% Indels: 57
DB: 1 Gaps: 10

US-09-208-629F-6 (1-408) x US-08-383-750-3 (1-1643)
```

```
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db 103 TATGCACATCACAGCGCCAGGATAGTAATAGCGCTCTGCATTACAGCGCTCGTCTTCATC 162
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIle 145
Db 163 ATTGGGCTCGTGGAAACTTACTAGCCTTGGTGGTCTGTTTCAAAACAGGAAATAATC 222
Qy 146 ---CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThr 164
Db 223 AACTCTACCACTCTATTCAACAATTTGGTGAATTTCTGATATATCTTTTACCACGGCT 282
Qy 165 LeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeu 184
Db 283 TTGCTTACAGTAATAGCTTACTATGCAATGGGCTTTGACTGGAGAAATCGAGATGCTTGT 342
Qy 185 CysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAla 204
Db 343 TGAGGAAATACGCGCTAGTGTGTTTACATCAACATATGCGAGGTGAACTTTATGACC 402
Qy 205 CysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuPro 224
Db 403 TGCTGAGTATTACCGCTTCTGCTGTTGGTGGACCTCTAGCTTACAAACAAGATAAAA 462
Qy 225 LysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMet 244
Db 463 AGGATTGAACATCAAAAGCGGTGTCATATTTGCTGGATTCTAGTATTGCTCAGACA 522
Qy 245 LeuProPhePheIle-----LeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
Db 523 CTCCTCATCTCATCAACCTTATGTCAAGAGGAG-----GCTGAAGG 567
Qy 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerSer----PropheGlnLeuTyr 280
```

```

Db      |||||      |||      |||      |||      |||      |||      |||
568 ATTACATGATGAGTATCCAAACTTTGAAGAACTAAATCTTCCCTGGATCTG--- 624

QY      |||      |||      |||      |||      |||      |||      |||
281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db      :|:      :|:      :|:      :|:      :|:      :|:      :|:
625 -----CTTGGGCGATGTTTCATAGGATATGTACTTCCACTTATATCAATCTCATCTGC 678

QY      |||      |||      |||      |||      |||      |||      |||
301 TyrAlaAlaIleIle----- 305

Db      |||      |||      |||      |||      |||      |||      |||
679 TATCTCAGATCTGTGCAAACTCTTCAGAACTGCCAAACAAACCCACTCACTGAGAAA 738

QY      |||      |||      |||      |||      |||      |||      |||
306 -----ArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320

Db      |||      |||      |||      |||      |||      |||      |||
739 TCTGGTGTAAACAAAAGGCTCTCAACAAATT----- 771

QY      |||      |||      |||      |||      |||      |||      |||
321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
Db      :|:      :|:      :|:      :|:      :|:      :|:      :|:
772 -----ATTCTTATTATTGTTGTTGTTCTCTCTGTTTCACACCTTACCATGTTGCAATT 825

QY      |||      |||      |||      |||      |||      |||      |||
341 IleIleHis-----HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db      |||      |||      |||      |||      |||      |||      |||
826 ATTCACATATGATTAAAGAGCTTCGTTTCTCTAAATTCCTGGAATGTAGCCAAAGACAT 885

QY      |||      |||      |||      |||      |||      |||      |||
355 LeuTyrPheIleTyrLeu---IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
Db      :|:      :|:      :|:      :|:      :|:      :|:      :|:
886 TCGTTCAGATTTCCTGACACTTTACAGTATGCTGATGAACCTCAATTGCTGCATGGAC 945

QY      |||      |||      |||      |||      |||      |||      |||
374 ProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys 393
Db      :|:      :|:      :|:      :|:      :|:      :|:      :|:
946 CCTTTATCTACTCTCTTGCATGATAA-----GGTATAAGAGAAAG 987

QY      |||      |||      |||      |||      |||      |||      |||
394 *****AsnAspLeuArgGluGln 401
Db      |||      |||      |||      |||      |||      |||      |||
988 GTTATGAGGATGCTGAAACGGCAA 1011

```

Search completed: June 29, 2003, 10:17:56
Job time : 104.056 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:26:54 ; Search time 192.736 Seconds
(without alignments)
4767.217 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDGTQVIMK.....AYLTXXNDLREQQPSQRT 408

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgn2_1/USPto_spool/US09208629/runat_24062003_101625_12912/app_query.fasta_1.1166
-DB=N_Geneseq_101002 -QMT=fastcap -SUFFIX=ring -MINWATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09208629 @CGN 1.1 27%runat_24062003_101625_12912 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	2145	99.4	6546	21	Human adenosine re
2	2142	99.3	1224	19	Human protease-act
3	2142	99.3	1830	21	Human low adenosin
4	2142	99.3	1830	21	Human adenosine re
5	2142	99.3	6546	21	Human low adenosin
6	1947	90.3	1102	19	Human protease-act
7	1443.5	66.9	1224	19	Mouse protease-act
8	1443.5	66.9	2409	21	Human protease-act
9	959	44.5	1124	19	Mouse protease-act
10	606.5	28.1	2732	16	Murine C140 recept
11	606.5	28.1	2732	17	Murine C140 recept
12	600.5	27.8	1477	16	Murine C140 recept
13	600.5	27.8	1477	17	Murine C140 recept
14	586.5	27.2	1451	21	Human PAR-2 DNA
15	585.5	27.1	1255	16	Human C140 recepto
16	585.5	27.1	1255	17	Human C140 recepto
17	585.5	27.1	1289	21	Human protease act
18	572.5	26.5	1414	16	Human C140 recept
19	572.5	26.5	1414	17	Human C140 recept
20	531.5	24.6	1361	20	Mouse protease-act
21	531.5	24.3	4895	20	Human protease-act
22	523.5	24.3	4895	21	Human protease act
23	523.5	24.3	4925	22	Human PAR4 coding
24	523	24.2	1534	20	Human protease-act
25	523	24.2	1534	21	Human low adenosin
26	523	24.2	1534	21	Human adenosine re
27	523	24.2	3472	20	Human thrombin rec
28	523	24.2	3480	21	Human thrombin rec
29	519	24.1	1278	24	Human CDNA encodin
30	516.5	23.9	1764	15	Fragment of the hu
31	512	23.7	2910	18	Thr-GPA1 fusion ge
32	512	23.7	3182	21	Human low adenosin
33	512	23.7	3182	21	Human adenosine re
34	507.5	23.5	1209	24	Human CDNA encodin
35	504	23.4	3480	13	Human thrombin rec
36	503.5	23.3	1116	24	Human CDNA encodin
37	496.5	23.0	1312	18	Mouse thrombin rec
38	488	22.6	2588	19	Mouse G-protein co
39	486.5	22.6	1080	19	Mouse G-protein co
40	466	21.6	1955	21	CDNA encoding a hu
41	465	21.6	1192	21	CDNA encoding a hu
42	465	21.6	2137	22	Human nGPCR5 codin
43	461	21.4	1080	21	Human orphan G pro
44	461	21.4	1080	21	Human G protein co
45	460	21.3	1080	22	Human G-protein co

ALIGNMENTS

RESULT 1

AAA35311
ID AAA35311 standard; DNA; 6546 BP.
XX AAA35311;
AC
XX
XX
XX 28-JUL-2000 (first entry)
XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:185.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
XX phosphothioate; impaired respiration; inflammation; allergy;
XX allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
XX antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;
XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX OS Homo sapiens.
 XX PN W0200009525-A2.
 XX PD 24-FEB-2000..
 XX PF 03-AUG-1999; 99WO-US17712.
 XX PR 03-AUG-1998; 98US-0095212.
 XX PA (UYEC-) UNIV EAST CAROLINA.
 XX PI Nyce JW;
 XX DR WPI; 2000-205971/18.
 XX
 XX New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -
 XX
 XX Disclosure; Page 1331-1333; 1343pp; English.
 XX
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasise to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A'-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 XX SQ Sequence 6546 BP; 1603 A; 1692 C; 1500 G; 1750 T; 1 other;

Alignment Scores:

Pred. No.:	9.57e-188	Length:	6546
Score:	2145.00	Matches:	405
Percent Similarity:	99.26%	Conservative:	0
Best Local Similarity:	99.26%	Mismatches:	3
Query Match:	99.44%	Indels:	0
DB:	21	Gaps:	0

US-09-208-629F-6 (1-408) x AAA35311 (1-6546)

Qy 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
 Db 88 TGCTCCCATGATTTTACAGATTTCACACGTTTAAAGACCGGAGCTCAGGTCATCAAAATG 147
 Qy 21 LysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSer 40
 Db 148 AAAGCCCTCATCTTTGACGTGCTGGCTCTCTGTTGTCACCTTTTGTTCAGAGT 207
 Qy 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
 Db 208 GGCATGGAAATGATACAAACAACTTGGCAAAGCCAACTTACCCATTAAAGACCTTTCGT 267

Qy 61 GlyAlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
 Db 268 GGAGCTCCCCAAATTCITTTGAAGAGTTCCTCTTCTGCTTGGAGCGTGGACAGA 327
 Qy 81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
 Db 328 GCCACGATTACTGTAAATAATTAAAGTGCCTGAAGAAAGTGTTCACATCTCCATGTGAAA 387
 Qy 101 AsnAlaThrMetGlyTyrIleuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
 Db 388 AATGCTACCATGGGTACCTGACCACTCTTAAAGTACTAAACTGATACCTGCCATCTAC 447
 Qy 121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
 Db 448 CTCCTGGTGTGTGTAGTGGTGTCCGGCCAAATGCTGTGACCTGTGGAGTGTCTTCTTC 507
 Qy 141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
 Db 508 AGGACCAGATCCATCTGTACCACTGTATTCTACCAACCTGGCCATTGCAGATTTCTT 567
 Qy 161 PheCysValThrIleuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
 Db 568 TTTTGTGTACATGGCTTTAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATT 627
 Qy 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
 Db 628 GGAGAGTCTCTGTGCGGGCCACACAGTCATCTTCTATGGCAACATGATGCTGCTCAT 687
 Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
 Db 688 CTGCTCTTGGCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCTTTCACCTAC 747
 Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
 Db 748 CGGGGCTCTCCCAAGCACACCTATGCTTGGTAAACATGTGGAGTGTGGGCAACAGTT 807
 Qy 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
 Db 808 TTTCTATATATGCTGCCATTTTTCATCTGAAGCAGGAATATTAATCTTGTTCAGCCAG 867
 Qy 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
 Db 868 ATCACCACCTGCATGATGTTCAACAACCTTGCAGTCTCTCATCTCCCTTCCAACCTAT 927
 Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleLysTyrCys 300
 Db 928 TACTTTCATCTCTTGGCATCTTGGATTTCTTAATTTCCATTTTGTGCTTATCATCTAC 987
 Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
 Db 988 TATGAGCCATCATCCGGACCTTAATGATACGATCATAGATGGTGTGGTATGTTAAG 1047
 Qy 321 AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleLeuLeu 340
 Db 1048 GCGAGTCTCTCATCTCTGTGATTTTACCATTTGCTTTTGTCTCCAAGCAATATTATCTT 1107
 Qy 341 IleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
 Db 1108 ATTATTACCATGCTAACTACTACTACAACAACACTGATGGCTTATATTATATATATCTC 1167
 Qy 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
 Db 1168 ATAGCTTTGTGCTGGGTAGTCTTAATAGTGTGCTTAGATCCATCTCTTTATTTCTCATG 1227
 Qy 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
 Db 1228 TCAAAACCAGAAATCACTCCACTGCTTACCTTACAAATAGTGAATATGATCTTAGAGAA 1287
 Qy 401 GlnGlyGlnProSerGlnArgThr 408
 Db 1288 CAAGGACAGCCCATCAGAGAACG 1311

RESULT 2
AAV07374
ID AAV07374 standard; cDNA; 1224 BP.

XX AC AAV07374;
XX DE 12-OCT-1998 (first entry)
XX DE Human protease-activated receptor 3 (PAR3) cDNA.

XX KW Protease-activated receptor 3; PAR3; thrombin receptor; human;
XX KW G-protein coupled receptor; agonist; antagonist; thrombosis;
XX KW atherosclerosis; restenosis; inflammation; blood coagulation;
XX KW blood clotting; heart attack; stroke; wound healing;
XX KW adult respiratory distress syndrome; glomerulosclerosis; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT CDS 58..1182
XX FT /*tag= a

XX PN WO9818456-A1.

XX PD 07-MAY-1998.

XX XX 29-OCT-1997; 97WO-US19732.

XX XX 30-OCT-1996; 96US-0742440.

XX XX (REGC) UNIV CALIFORNIA.

XX XX Connolly A, Coughlin SR, Ishihara H;

XX DR WPI; 1998-271905/24.

XX DR P-PSDB; AAW51406.

XX PT DNA encoding protease-activated receptor 3 - for detection of
XX PT specific agonists and antagonists, potentially useful for treating
XX PT e.g. thrombosis, atherosclerosis, inflammation etc.

XX PS Claim 3; Page 40-41; 74pp; English.

XX CC This cDNA clone codes for mouse protease-activated receptor 3
XX CC (PAR3) (see AAW51406), a cell surface protein which is specifically
XX CC activated by thrombin or a thrombin agonist, thereby activating
XX CC signalling events such as phosphoinositide hydrolysis, calcium ion
XX CC efflux and platelet aggregation. A cDNA library was established
XX CC from rat platelets and amplified (primers given in AAV07376-77). A
XX CC product was identified that was expected to encode a new G-protein
XX CC coupled receptor, related to PAR1 or 2. This was used to isolate
XX CC human and murine DNA by a combination of PCR and hybridisation
XX CC techniques. Murine cDNA (see AAV07372), murine genomic DNA (see
XX CC AAV07373), human cDNA and human genomic DNA (see AAV07375) sequences
XX CC are provided, as well as the murine (see AAW51405) and human PAR3
XX CC proteins. Also claimed are vectors, host cells and an assay
XX CC to detect. Host cells are used to screen compounds for their ability
XX CC to act as agonists or antagonists of the effects of thrombin-PAR3
XX CC interaction. Agonists are used to treat wounds, thrombosis,
XX CC atherosclerosis, restenosis, inflammation and other thrombin
XX CC activated disorders. Antagonists (see AAW51415-21) are used to
XX CC control blood coagulation and thereby to treat heart attack and
XX CC stroke. They also mediate inflammatory and proliferative responses
XX CC to injury as occur in wound healing, atherosclerosis, restenosis,
XX CC pulmonary inflammation (ARDS) and glomerulosclerosis.

XX SQ Sequence 1224 BP; 303 A; 316 C; 221 G; 384 T; 0 other;

Alignment Scores:

Pred. No.: 2, 08e-188 Length: 1224
Score: 2142.00 Matches: 405
Percent Similarity: 99.26% Conservativity: 0
Best Local Similarity: 99.26% Mismatches: 3

Query Match:
DB: 99.30% Indels: 0
19 Gaps: 0

US-09-208-629F-6 (1-408) x AAV07374 (1-1224)

QY	1	CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet	20
DB	1	TGCTCCATGATTTTACAGATTTTCATACGTTTAAAGAGACGGGACTCAGGTCATCAAAATG	60
QY	21	LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer	40
DB	61	AAAGCCCTCATCTTTGCAGCTGCTGCCCTCTCTGTTGTCCTCAGTCTTTTGTCTCAGAGT	120
QY	41	GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrIleuProIleLysThrPheArg	60
DB	121	GGCATGGAAATGATACAAACAACTTGGCAAGCCAACTTACCCATTAAGACCTTTCGT	180
QY	61	GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly	80
DB	181	GGAGCTCCCCCAAAATTTCTTTGAAGAGATTTCCCTTTTCTGCCCTTGAAGGCTGGACAGGA	240
QY	81	AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys	100
DB	241	GCCACGATTACTGTAAAAAATTAAAGTCCCTGAAAGAAAGTGCTTCACATCTCCATGTGAAA	300
QY	101	AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr	120
DB	301	AATGCTACCATGGGTACCTGACCAGCTCTTAAAGTACTAAACCTGATACCTGCTCATCTAC	360
QY	121	LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe	140
DB	361	CTCCTGGTGTGTAGTGTGGTGTCCCGGCCAAATGCTGTGACCTGTGGATGCTTTTCTTC	420
QY	141	ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu	160
DB	421	AGGACCATTCATCTGTACCATCTGTATCTACACCAACCTGGCCATTTGAGATTTCTT	480
QY	161	PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe	180
DB	481	TTTTGTGTACATTTGCCCTTTTAAAGTAGCTTATCATCTCAATGGGAACAACCTGGGTATTT	540
QY	181	GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle	200
DB	541	GGAGAGTCTCTGTCCGGGCCACACAGTCATCTTCTATGGCAACATGTACTCTCCATT	600
QY	201	LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr	220
DB	601	CTGCTCTTGGCTTGCATCAGCATCAACCGCTACTGGCCATCTGCTCATCTTTCACCTAC	660
QY	221	ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal	240
DB	661	CGGGGCTGCCCAAGCACACCTATGCTTGGTAACATGTGGACTGGTGTGGCAACAGATT	720
QY	241	PheLeuTyrMetLeuProPhePheLeuLeuLysGlnGluTyrTyrLeuValGlnProAsp	260
DB	721	TTCTTATATATGCTGGCAATTTTTCATCTACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC	780
QY	261	IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr	280
DB	781	ATCACCACCTGGCCATGATGTTTCAACACATTTGGAGTCCCTCATCTCCCTTCCAACTCTAT	840
QY	281	TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys	300
DB	841	TACTTCATCTCTTGGCATCTTTTGGATCTTAAATTCATCTTGTGCTTATCATCTACTGC	900
QY	301	TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys	320
DB	901	TATGCAGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTGTGTGTATGTTAAG	960
QY	321	AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleLeu	340
DB	961	GCAGAGTCTCTCTATCTTGTGATTTTACCATTGTCTTTGCTTTCCTCAAGCAATATTATCTT	1020

QY 341 IleIleHisAlaAenTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
 DB 1021 ATATTACCATGCTAACTACTACTACAAACACATGATGGCTATATTTATATATC 1080
 QY 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
 DB 1081 ATAGCTTGTGCTGGGTAGTCTTAATAGTGTGCTTAGATCCATTCCTTTATTTCTCATG 1140
 QY 381 SerIysThrArgAnHisSerThrAlaTyrLeuThrIys*****AsnAspLeuArgGlu 400
 DB 1141 TCAAAACCCAGAAATCACTCCACTGCTTACCTTACAAATAGTGAATGATCTTAGAGAA 1200
 QY 401 GlnGlyGlnProSerGlnArgThr 408
 DB 1201 CAGGACAGCATCACAGAGACG 1224

RESULT 3

AAF21430
 ID AAF21430 standard; DNA; 1830 BP.

XX AC AAF21430;
 XX AC AAF21430;
 DT 14-MAR-2001 (first entry)
 XX DE Human low adenosine antisense oligonucleotide related sequence #2997.
 XX KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.

OS Homo sapiens.

XX W0200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

PA (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -

XX Disclosure; Page 1414-1415; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and/or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central

CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.

XX SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;

Alignment Scores:
 Pred. No.: 3 49e-188 Length: 1830
 Score: 2142.00 Matches: 405
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 21 Gaps: 0

US-09-208-629F-6 (1-408) x AAF21430 (1-1830)

QY 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
 DB 88 TGCTCCATGATTTTACAGATTTTACATACGTTTAAAGACGGGACTCAGGTCACTCAAAATG 147
 QY 21 LysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSer 40
 DB 148 AAGCCCTCATCTTTGACGTGCTGCTGCTCTCTGTTGCTGCTCTCTGTTGCTGCTGCT 207
 QY 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
 DB 208 GGCTGGAAATGATACAAACAACTTGGCAAGCCAACTTACCCTTAAGACCTTTCGT 267
 QY 61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTyrThrGly 80
 DB 268 GGAGCTCCCCAAATTTCTTTGAAGAGTTCCCTTTTCTGCTTGAAGCTGGACAGGA 327
 QY 81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
 DB 328 GCCACGATTACTGTAAAAATTAAGTCCCTCGAAGAAAGTCTTCACATCTCCATGTGAAA 387
 QY 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
 DB 388 AATGCTACCATGGGTACCTGACCACTCTTAACTAACTGATACCTGCACTAC 447
 QY 121 LeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
 DB 448 CTCCTGGTGTGTAGTTGGTGTCCCGGCAATGTCTGACCTGTGGATGCTTTCTTC 507
 QY 141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
 DB 508 AGGACAGATCCATCTGTACCTGTATTCTACCAACCTTGCCCACTTGAGATTTTCTT 567
 QY 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
 DB 568 TTTTGTGTACATGGCTTTTAAAGATAGCTTATCTCATCTCAATGGGCAACACTGGGTATT 627
 QY 181 GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
 DB 628 GGAGAGTCTCTGTGCGGGCCACACAGTCATCTTCTATGGCAACATGTACTGCTCCATT 687
 QY 201 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
 DB 688 CTGCTCTTGGCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCTCTTACACCTAC 747


```
Qy 221 ArgGlyLeuProLysHisThrTyAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 748 CGGGCCCTGCCCAAGCACACATATGCTTGGTAACATGTCAGCTGGTGGCAACAGTT 807
Qy 241 PheLeuTyrMetLeuProPhePheLeuLeuGlnGluTyrLeuValGlnProAsp 260
Db 808 TTCCTATATATGCTGCCCATTTTCATCTACTGAAGCAGGAATATATCTGTGTGAGCCAGAC 867
Qy 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
Db 868 ATCACCACCTGCCATGATGTTCAACACATTCGAGTCCCTCATCTCCCTCCCACTCTAT 927
Qy 281 TyrPheLeuSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuTyrCys 300
Db 928 TACTTTCATCTCCTTGGCATTTCTTGGATTTCTTAATTCATTTGTCATTATCATCTACTGC 987
Qy 301 TyrAlaAlaLeuLeuLeuLeuValLeuValLeuValLeuValLeuValLeuValLeu 320
Db 988 TATGCACCATCATCCGACACTTAATGTCATACCATCATGATGTTGTGTATGTTAAG 1047
Qy 321 AlaSerLeuLeuLeuLeuValLeuValLeuValLeuValLeuValLeuValLeuVal 340
Db 1048 GCGAGTCTCTCATCTCTGTTGATTTTACCATTTGCTTGGCTCCAAACATATATATCTT 1107
Qy 341 IleLeuHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheLeuValLeu 360
Db 1108 ATTATTCACCATGTAACTACTACTACAAACAACTGATGCTTATATTTATATATATCTC 1167
Qy 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
Db 1168 ATAGCTTTGTCCTGGGTAGTCTTAATAGTTGCTTAGATCCATTCCTTTATTTCTCATG 1227
Qy 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys****AsnAspLeuArgGlu 400
Db 1228 TCAAAACCCAGAAATCATCTCACTGCTTACCTTACAAATAGTCAATGATCTTAGAGAA 1287
Qy 401 GlnGlyGlnProSerGlnArgThr 408
Db 1288 CAAGGACAGCCATCACAGAGAACG 1311
RESULT 4
AAA35308
ID AAA35308 standard; DNA; 1830 BP.
XX
AC AAA35308;
XX
28-JUL-2000 (first entry)
DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:182.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
```

DR WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers

XX Disclosure; Page 1329; 1343pp; English.

XX The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1880
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

SQ Sequence 1830 BP; 473 A; 464 C; 337 G; 556 T; 0 other;

Alignment Scores:
Pred. No.: 3,49e-188 Length: 1830
Score: 2142.00 Matches: 405
Percent Similarity: 99.26% Conservatives: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 21 Gaps: 0

US-09-208-629F-6 (1-408) x AAA35308 (1-1830)

Qy 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
Db 88 TGCTCCATGATTTTACAGATTTTCATACCGTTTAAGAGACGGGACTCAGGTTCATCAAAATG 147
Qy 21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
Db 148 AAAGCCCTCATCTTTGCAGCTGCTGCGCCCTCTCTCTGTGTGCCACTTTTGTGACAGT 207
Qy 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
Db 208 GGCATGGAAATCATACAAACAACTTGGCAACCAACCTTACCCATTAGACCTTTCGT 267
Qy 61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGly 80
Db 268 GGAGCTCCCCCAATTTCTTTGAAGAGTTCCTCTTTTCTGCTTGAAGCTGGACAGGA 327
Qy 81 AlaThrIleThrValLysIleLysCysProGluSerAlaSerHisLeuHisValLys 100
Db 328 GCCACGATTACTGTAATAAATAAGTCCCTGGAAGAAGTCTTCACATCTCCATGTGAAA 387
Qy 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
Db 388 AATGCTACCATGGGGTACCTGACCACTCTTAAGTACTAACTGATACCTGCATCTAC 447
Qy 121 LeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140

Db	448	CTCTGTGTGTTGTAGTGTGGTGTCCTCCGGCCAAATGCTGTGACCCCTGTGGATGCTTTTCTTC	507
Qy	141	ArgThrArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu	160
Db	508	AGGACCAAGATCCATCTGTACCACTGTATTCTACACCAACCTGGCCATTGGCAGATTTTCTT	567
Qy	161	PheCysValThrLeuProPheLeysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe	180
Db	568	TTTTTGTGTACATTGGCCCTTTAAGATAGCTTATCATCTCAATGGAAACAACTGGGTATTT	627
Qy	181	GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle	200
Db	628	GGAGAGTCTGTGTCGGGGCCACACAGTCATCTTCTATGGCAACATGTAAGTCTCCATT	687
Qy	201	LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr	220
Db	688	CTGCTCTTGTGCTGCATCAGCATCAACCGCTACCTGGCCATCGTCCATCTCTTCACCCATC	747
Qy	221	ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal	240
Db	748	CGGGGCTGGCCCAAGCACACCTATGCTCTGGTAACATGTGCATGGTGGGCAACAGTT	807
Qy	241	PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp	260
Db	808	TTCTTATATATGCTGCCATTTTTCATACTGAAGCAGGAATATTATCTTGTTCAGCCAGAC	867
Qy	261	IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr	280
Db	868	ATCACCACTGCCATGATGTTTCAACAACATGGCAGTCCCTCATCTCCCTTCCAACCTCAT	927
Qy	281	TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys	300
Db	928	TACTTTCATCTCTTGGCATCTTTGGATTCTTAATTCATTTGTGCTTATCATCTACTGCG	987
Qy	301	TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgThrLeuTrpTyrValLys	320
Db	988	TATCAGCCATCATCCGGACACTTAAATGCATACATCATAGATGGTTGTGGTATGTTAAG	1047
Qy	321	AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu	340
Db	1048	CGAGATCTCTCATCTTGTGGATTTTACCATTGCTTTGCTCCAGCAATATTATTCTT	1107
Qy	341	IleIleHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu	360
Db	1108	ATTATTACCATTGCTAACTACTACTACAAACACTGATGGCTTATATTATATATATCTC	1167
Qy	361	IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet	380
Db	1168	ATAGCTTTGTGCCCTGGGGAGTCTTAAATAGTTGCTTAGATCCATTCCCTTTATTTTCTCATG	1227
Qy	381	SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu	400
Db	1228	TCAAAACACAGAAATCACTCCACTGCTTACCTTACAAATAGTGAATGATCTTAGAGAA	1287
Qy	401	GlnGlyGlnProSerGlnArgThr	408
Db	1288	CAAGCAGACCCATCACAGAGAACG	1311

RESULTS

RESULT 5
AAE21433

AAF21433
ID AAF21433 standard: DNA: 6546 BP.

ID
XX
AAE21433

XX AAF21433:

AC
XX
AAF21433;

DT: 14-MAR-2001 (first entry)

DI. 14-MAR-2001 (FIRST ENTRY)
XX

Human low adenosine antiserum

DE
HUMAN LOW AEROBIC
XX

KW Low adenosine antisense ol

KW human; airway disorder; bronchitis

KW surfactant depletion; resp

KW immunosuppressive; antiast

KW respiratory obstruction; p

surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

WO2000062736-A2

26-OCT-2000

24-MAR-2000: 2000WO-IJS08020

06-APR-1999: 99IIS-0127958

(IVEC-) INTV EAST CAROLINA

(NYCE/) NYCE J W:

Nyce JW;

WPI; 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

Disclosure; Page 1416-1418; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base oligonucleotide, (II) can have respiratory, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF19434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 6546 BP: 1603 A: 1691 C: 1500 G: 1751 T; 1 other;

Alignment Scores:		
Pred. No.:	1.81e-197	Length:
Score:	2142.00	Matches:
Percent Similarity:	99.26%	Conservative:
Best Local Similarity:	99.26%	Mismatches:
Query Match:	99.30%	Indels:
DB:	21	Gaps:
		6546

US-09-208-629F-6 (1-408) x AAF21433 (1-6546)

Qy 1 CysSerMetIleLeuGlnIleSer***ArgIleuArgAspGlyThrGlnValIleIysMet 20

Db 88 TGCTCCATGATTTTACAGATTTTCATAACGTTTAAAGACGGGACTCAGGTCACTCAAAATG 147
Qy 21 LysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSer 40
Db 148 AAAGCCCTCATCTTTGTCAGCTGCTGCTCTCTGCTTCTGTTGCCACTTTTGTGCAGAGT 207
Qy 41 GlyMetGlnAsnAspThrAsnAsnLeuAlaAlaPheProThrLeuProIleLeuThrPheArg 60
Db 208 GGCATGGAAATGATACAAACACTTGGCAAGCCAACTTACCCATTAAGACCTTTTCGT 267
Qy 61 GlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTTPThrGly 80
Db 268 GAGCTCCCCCAAAATCTTTTGAAGAGTTCCCTTTTCTGCTTGGAGGCTGGACAGGA 327
Qy 81 AlaThrIleThrValIleLysCysProGluSerAlaSerHisLeuHisValLys 100
Db 328 GCCACGATTACTGTAAAAATTAAGTGCCCTGAAGAAAGTCTTCACATCTCCATGTGAAA 387
Qy 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
Db 388 AATGCTACCATGGGGTACCTGACCAGCTCTCTTAAGTACTAACTGATPACCTGGCATCTAC 447
Qy 121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTTPMetLeuPhePhe 140
Db 448 CTCCTGTGTGTGTAGTGTGTGTCGCCGCAATGCTGTGACCTGTGGATGCTTTTCTTC 507
Qy 141 ArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
Db 508 AGGACACAGATCCATCTGACCACTGTATTCTACACCAACCTGGCCATTTGAGATTTCCT 567
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db 568 TTTTGTGTTACATTGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATTT 627
Qy 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 628 GGAAGAGTCTGTGCGGGCCACACAGTATCTTCTATGGCAACATGTACTGCTCCATT 687
Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db 688 CTGCTCTTGGCTGCATCAGCATCAACCGCTACCTGGCCATCTGCCATCTTTCACTAC 747
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTTPAlaThrVal 240
Db 748 CGGGGCTGCGCAAGCACACTATGCTTGGTAACATGTGACTGCTGGTGGCAACAGTT 807
Qy 241 PheLeuTyrMetLeuProPheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db 808 TTCTTATATATGCTGCCATTTTTCATATCTGAAGCAGGAATATTATCTTGTTCAGCCAGAC 867
Qy 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
Db 868 ATCACCACCTGCCATGATGTTTCAACAACATGTCGAGTCTCACTCCCTTCCAACTCTAT 927
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 928 TACTTCACTCTCTGGCATCTTTGGATTTCTTAATTCATTGTTGCTTATCATCTACTGC 987
Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTTPLeuTTPThrValLys 320
Db 988 TATGACGCCATCATCCGGACACTTAAATGTCATACATCATAGATGGTGTGGTATGTTAAG 1047
Qy 321 AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
Db 1048 GCGAGTCTCTCATCTCTGTGATTTTACCATTTGCTTGGTCCAGCAATATATTCTT 1107
Qy 341 IleIleHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
Db 1108 ATTATTCCATGTCTAACTACTACTACAACAACACTGATGGCTTATATTTATATATCTC 1167
Qy 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
Db 1168 ATAGCTTGTGCGGTGAGTCTTAAATAGTTGCTTAGATCCATCTCTTTATTTTCTCATG 1227

Qy 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1228 TCAAAACCCAGAAATCACTCCACTGCTTACCTTACAAATAGTGAATGATCTTAGAGAA 1287

Qy 401 GlnGlyGlnProSerGlnArgThr 408
Db 1288 CAAGGACGCCATCACAGAGAAG 1311

RESULT 6

AAV07375

ID AAV07375 standard; DNA; 1102 BP.

XX AAV07375;

AC AAV07375;

XX 12-OCT-1998 (first entry)

DT 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

XX 12-OCT-1998 (first entry)

SQ Sequence 1102 BP; 276 A; 286 C; 195 G; 345 T; 0 other;

Alignment Scores:

Pred. No.: 1.86e-170 Length: 1102
 Score: 1947.00 Matches: 364
 Percent Similarity: 99.46% Conservative: 1
 Best Local Similarity: 99.18% Mismatches: 2
 Query Match: 90.26% Indels: 0
 DB: 19 Gaps: 0

US-09-208-629f-6 (1-408) x AAV07375 (1-1102)

Qy 40 SerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProLysThrPhe 59
 Db 1 ACAGGATGAAATATGATACAAACAACTTGGCAAGCCAACTTACCCATTAAGACCTTT 60
 Qy 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTTPThr 79
 Db 61 CGTGGAGCTCCCAAAATCTTTTGAAGAGTTCCCTTTTCTGCTTGGAGGCTGACA 120
 Qy 80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
 Db 121 GGAGCCACGATTACTGTAAAAATTAAAGTCCCTGAGAAAGTCTTCACATCTCCATGTG 180
 Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLleProAlaIle 119
 Db 181 AAAAATGCTACCATGGGTACCTGACCACTCTTAAGTACTAACTGATACCTGCCATC 240
 Qy 120 TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTTPMetLeuPhe 139
 Db 241 TACCTCTGGTGTGTAGTTGGTCCCGGCATGCTGTGACCTCGTGGATGCTTTTC 300
 Qy 140 PheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPhe 159
 Db 301 TTCAGGACCAAGATCCATCTGACACATGTTCTACACCAACCTGGCCATGTCAGATTTT 360
 Qy 160 LeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTTPVal 179
 Db 361 CTTTTTGTGTGTACATGGCTTTTAAAGATAGCTTATCATCTCAATGGGAACAACATGGGTA 420
 Qy 180 PheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer 199
 Db 421 TTGGAGAGGTCCTGTGCGGGGCCACCAAGTATCTTATGGCAACATGTACTGCTCC 480
 Qy 200 IleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
 Db 481 ATTCGTCTCTGCTGCATCAGATCAACGGCTACCTGGCCATCGTCCATCTTTCAACC 540
 Qy 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTTPAlaThr 239
 Db 541 TACCGGGCCCTGCCCAAGCACACCTATGCTTGGTAACATGTGACTGGTGGGCAACA 600
 Qy 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGluTyrTyrLeuValGlnPro 259
 Db 601 GTTTTCTTATATGCTGCCATTTTTCATCTGAGCAGGAATATATCTTGTTCAGCCA 660
 Qy 260 AspileThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeu 279
 Db 661 GACATCACCACTGCCATGATGTTTCAACACTTGGCAGTCTCATCTCCCTCCAACTC 720
 Qy 280 TyrTyrPheLeuSerLeuAlaPhePheGlyPheLeuLeuProPheValLeuIleTyr 299
 Db 721 TATTACTTCATCTCTTGGCATCTTTTGGATTTCTTAATTCATATGCTTATCATCTAC 780
 Qy 300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTTPLeuTTPTyrVal 319
 Db 781 TGCTATGAGCCATCATCCGAGACACTTAATGTCATACCATCATAGATGTTGTGTATGTT 840
 Qy 320 LysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIle 339
 Db 841 AAGCGAGTCTCTCATCTTGTGATTTTACCATTTGCTTGTGCTCCAGCAATATATT 900
 Qy 340 LeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyr 359
 Db 901 CTTATTATTACCATGCTAACTACTACTACTACAAACACATGATGCTTATATTTATATAT 960

Qy 360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
 Db 961 CTCATAGCTTTTGGCTGGTAGTCTTAATAGTCTTGTAGATCCATTCCTTTATTTCTC 1020
 Qy 380 MetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArg 399
 Db 1021 ATGTCAAAACCAGAAATCACTCCACTGCTTACCTTACAAATAGTGAATATGATCTTAGA 1080
 Qy 400 GluGlnGlyGlnProSerGln 406
 Db 1081 GAACAAGGACAGCCATCAG 1101
 RESULT 7
 AAV07372
 ID AAV07372 standard; cDNA; 1224 BP.
 XX AAV07372;
 AC AAV07372;
 DT 12-OCT-1998 (first entry)
 XX Mouse protease-activated receptor 3 (PAR3) cDNA.
 DE
 XX Protease-activated receptor 3; PAR3; thrombin receptor; mouse;
 KW G-protein coupled receptor; agonist; antagonist; thrombosis;
 KW atherosclerosis; restenosis; inflammation; blood coagulation;
 KW blood clotting; heart attack; stroke; wound healing;
 KW adult respiratory distress syndrome; glomerulosclerosis; ds.
 XX
 OS Mus sp.
 XX
 XX Key Location/Qualifiers
 FH 51..1160
 FT CDS /*tag= a
 FT
 XX MO9818456-A1.
 PN
 XX 07-MAY-1998.
 PD
 XX 29-OCT-1997; 97WO-US19732.
 PF
 XX 30-OCT-1996; 96US-0742440.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Connolly A, Coughlin SR, Ishihara H;
 PI
 XX WPI; 1998-271905/24.
 DR P-PSDB; AAW51405.
 XX
 PT DNA encoding protease-activated receptor 3 - for detection of
 PT specific agonists and antagonists, potentially useful for treating
 PT e.g. thrombosis, atherosclerosis, inflammation etc.
 XX
 PS Claim 3; Page 38-39; 74pp; English.
 XX
 CC This cDNA clone codes for mouse protease-activated receptor 3
 CC (PAR3) (see AAW51405), a cell surface protein which is specifically
 CC activated by thrombin or a thrombin agonist, thereby activating
 CC signalling events such as phosphoinositide hydrolysis, calcium ion
 CC efflux and platelet aggregation. A cDNA library was established
 CC from rat platelets and amplified (primers given in AAV07376-77). A
 CC product was identified that was expected to encode a new G-protein
 CC coupled receptor, related to PAR1 or 2. This was used to isolate
 CC human and murine DNA by a combination of PCR and hybridisation
 CC techniques. Murine cDNA, murine genomic DNA (see AAV07373), human
 CC cDNA (see AAV07374) and human genomic DNA (see AAW51406) PAR3
 CC provided, as well as the murine and human (see AAW51406) PAR3
 CC proteins. Also claimed are vectors, host cells and an assay
 CC device. Host cells are used to screen compounds for their ability
 CC to act as agonists or antagonists of the effects of thrombin-PAR3
 CC interaction. Agonists are used to treat wounds, thrombosis,
 CC atherosclerosis, restenosis, inflammation and other thrombin
 CC activated disorders. Antagonists (see AAW51415-21) are used to

CC control blood coagulation and thereby to treat heart attack and
 CC stroke. They also mediate inflammatory and proliferative responses
 CC to injury as occur in wound healing, atherosclerosis, restenosis,
 CC pulmonary inflammation (ARDS) and glomerulosclerosis.
 XX
 SQ Sequence 1224 BP; 299 A; 341 C; 239 G; 345 T; 0 other;

Alignment Scores:

Pred. No.: 6.87e-124 Length: 1224
 Score: 1443.50 Matches: 264
 Percent Similarity: 83.74% Conservatives: 45
 Best Local Similarity: 71.54% Mismatches: 59
 Query Match: 66.92% Indels: 1
 DB: 19 Gaps: 1

US-09-208-629F-6 (1-408) x AAV07372 (1-1224)

Qy 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
 Db 33 GGGTCTCAGGACATCAAGATGAATAATCTTATCTTGGTTGAGCTGGGCTGTTCTG 92
 Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThr 53
 Db 93 CCAGTCACTGTTGCCAAGTGGGATA--AATGTTTCAGACAACTCAGCAAGCCCAACC 149
 Qy 54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73
 Db 150 TTAACATTAAGAGCTTTAATGGGGTCCCAAAATACCTTGAAGAATTTCCCACTTCT 209
 Qy 74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
 Db 210 GACATAGAGGGCTGGACAGGAGCCACCAACTATATAAGCGGAGTGTCCGAGGACGT 269
 Qy 94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTrpLeuThrSerSerLeuSerThr 113
 Db 270 ATTTCAACTTCCACGTAATATGCTACCATAGATGACCTTGAGAGTTCCTTAAGTACC 329
 Qy 114 LysLeuIleProAlaIleLysLeuLeuValPheValGlyValProAlaAsnAlaVal 133
 Db 330 CAAGTGATACCTGCATCTATATCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 389
 Qy 134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheThrAsn 153
 Db 390 ACCCTGTGGAACTCTCCTTAAGACCAATCCATCATGCTGTGTCATCTTCCACCAAC 449
 Qy 154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaThrHisLeu 173
 Db 450 CTGGCCATCGCAGATCTCCTTTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509
 Qy 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheThr 193
 Db 510 AATGGCAACAACCTGGGTATTTGGCGAGGTGATGTCGGGATCACCACGGTGTCTTCTAC 569
 Qy 194 GlyAsnMetTrpCysSerIleLeuLeuAlaCysIleSerIleAsnArgThrLeuAla 213
 Db 570 GGCAACATGACTGCGTATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 629
 Qy 214 IleValHisProPheThrTrpArgGlyLeuProLysHisThrTrpAlaLeuValThrCys 233
 Db 630 ACGGCTCACCTTTCACATACCAAGAGTGCCTGCAAGCGAGTCTCTTCTGCTCATGTGT 699
 Qy 234 GlyLeuValTrpAlaThrValPheLeuTrpMetLeuProPhePheIleLeuLysGlnGlu 253
 Db 690 GGCATAGTGGGTGATGTTTCTTATACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 749
 Qy 254 TyrTrpLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
 Db 750 TACCACCTCGTCCACTCAGAGATCACCACCTGCCACGATGCTGCTGCTGCTGCTGCTGCT 809
 Qy 274 SerSerProPheGlnLeuTrpPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
 Db 810 CCATCATCTTCGATTTCTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 869

Qy 294 PheValLeuIleLysTyrCysTyrAlaAlaIleAlaIleArgThrLeuAsnAlaTyrAspHis 313
 Db 870 TTTGTGATCATCATCT 929
 Qy 314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThrIleCysPhe 333
 Db 930 ATATGGCTGGGTACATCAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 989
 Qy 334 AlaProSerAsnIleIleLeuIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353
 Db 990 GCCCCCAACCAACATCATCTCTGTAATCCACCACTGCACTACTACTACCACATACCGAC 1049
 Qy 354 GlyLeuTyrPheIleTyrIleLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373
 Db 1050 AGCTTGTACTTATGATCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1109
 Qy 374 ProPheLeuTyrPheLeuMetSerLys 382
 Db 1110 CCATTCCTTACTTTGTGATGCGAAA 1136

RESULT 8

AAZ50774
 ID AAZ50774 standard; DNA; 2409 BP.
 XX AAZ50774;
 XX 31-MAY-2000 (first entry)
 XX Human protease activated receptor-3 DNA.
 XX Human; PAR-3; protease activated receptor;
 KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell; ds.
 OS Homo sapiens.
 XX WO200008150-A1.
 PD 17-FEB-2000.
 XX 05-FEB-1999; 99WO-IL00079.
 PF 07-AUG-1998; 98IL-0125698.
 PR (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 PA Bar-Shavit R;
 PI WPI; 2000-205706/18.
 DR

XX Treating metastatic tumor cells useful for treating disorders involving
 PT placenta implantation in a female comprises administration of an
 PT antisenase molecule complementary to an RNA sequence of a protease
 XX activated receptor protein -
 XX Example 3; Fig 10; 46pp; English.

CC The patent discloses a method to treat metastatic tumour cells using
 CC an antisenase molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisenase molecules and
 CC antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a human PAR-3
 CC DNA used for producing antisenase molecules for treating invasive
 CC cells.

SQ Sequence 2409 BP; 598 A; 652 C; 454 G; 703 T; 2 other;

Alignment Scores:

Pred. No.: 1.65e-123 Length: 2409
 Score: 1443.50 Matches: 264
 Percent Similarity: 83.74% Conservatives: 45
 Best Local Similarity: 71.54% Mismatches: 59

Query Match:	66.92%	Indels:	1
DB:	21	Gaps:	1
US-09-208-629f-6 (1-408) x AA250774 (1-2409)			
Qy	14	GlyThrGlnValIleLeuValMetLeuAlaLeuPheAlaAlaLeuLeuLeuLeu 33	
Db	135	GGGTCTCAGGACATCAAGATGAAATCCTTATCTTGGTTGAGTGGCTGCTGTTCTG 194	
Qy	34	LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLeuProThr 53	
Db	195	CCAGTCACTGTTTGGCCAAAGTGCATA---AATGTTTCAGACCAACTCAGCAAGCAAC 251	
Qy	54	LeuProIleLeuThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73	
Db	252	TTAACATTAAGAGTTTAAATGGGGTCCCAAAATACCTTTGAAGAAATCCCACTTCT 311	
Qy	74	AlaLeuGluGlyTrpThrGlyAlaThrIleThrValIleLeuValCysProGluGluSer 93	
Db	312	GACATAGAGGCTGGACAGGAGCCACCAACTATAAAGCGAGTGTCCCGAGGACAGT 371	
Qy	94	AlaSerHisLeuHisValIleValAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113	
Db	372	ATTTCACACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAAGTTCCTTAAGTACC 431	
Qy	114	LysLeuIleProAlaIleTyrLeuValPheValGlyValProAlaAsnAlaVal 133	
Db	432	CAAGTATACCTGCGCATCTATCTCTGCTGTTGTGGTTGGTGTACCATCAACATCGTG 491	
Qy	134	ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153	
Db	492	ACCTGTGGAAACTCTCTTAAGGACCAATCCATCAGTCTGGTCACTCTTTACACCAAC 551	
Qy	154	LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheIleAlaTyrHisLeu 173	
Db	552	CTGGCCATCGCAGATCTCTCTTTCTGTGCACACTGCCATTTAAGATCGCCTACCATCTC 611	
Qy	174	AsnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyr 193	
Db	612	ANTGGCAACAACTGGGTATTGGCGAGTGTATGGCCGGATCACCAGTGTCTTCTTCAT 671	
Qy	194	GlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213	
Db	672	GGCAACATGTACTGCGTATCTCTGATCTCACTTGCATGGCATCAACCGCTACCTGGCC 731	
Qy	214	IleValHisProPheThrThrArgGlyLeuProIleHisThrTyrAlaLeuValThrCys 233	
Db	732	ACGGCTCACCCCTTTACATACCAAGAGCTGCCAAACGCGAGTCTCTCTTGTCTCATGTGT 791	
Qy	234	GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuValGlnGlu 253	
Db	792	GGCATAGTGGGTGATGGTTTCTTATACATGCTGCCCTTTGTATCTCTGAGACAGAG 851	
Qy	254	TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273	
Db	852	TACACCTCGTCCACTCAGATACACACCTGCCACCATGTCTGTCGAGCGGTGGAGTCC 911	
Qy	274	SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuLeuPro 293	
Db	912	CCATCATCTCTCCGATTTACTACTCTCTCTCTCTTAGCATTTCTTGGGTTCTCTCATCCG 971	
Qy	294	PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313	
Db	972	TTTGTGATCATCTCTCTGTATACAGACTCTCTACCAAACTTAATCAAGGATCGG 1031	
Qy	314	ArgTrpLeuTrpTyrValIleValAsnLeuLeuValIlePheThrIleCysPhe 333	
Db	1032	ATATGGTGGGTACATCAAGGCGTCTCTCTCATCTCTGTGATTTTCACAATTTGCTTT 1091	
Qy	334	AlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353	
Db	1092	GCCCCCAACCAATCATACTACTCGTAATCCCACTGCACTACTACTACCAATACCGAC 1151	
Qy	354	GlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAsp 373	
Db	1152	AGCTTGACTTTTATGATCTTATTTGCTGTGGGAGCCCTGAATAGCTGCCTAGAT 1211	
Qy	374	ProPheLeuTyrPheLeuMetSerIlys 382	
Db	1212	CCATTCCTTTACTTTGTCTGATGCGAA 1238	
RESULT 9			
ID	AAV07373	standard; DNA; 1124 BP.	
XX	AAV07373;		
XX	12-OCT-1998	(first entry)	
XX	Mouse protease-activated receptor 3 (PAR3) genomic DNA.		
XX	Protease-activated receptor 3; PAR3; thrombin receptor; mouse;		
KW	G-protein coupled receptor; agonist; antagonist; thrombosis;		
KW	atherosclerosis; restenosis; inflammation; blood coagulation;		
KW	blood clotting; heart attack; stroke; wound healing;		
KW	adult respiratory distress syndrome; glomerulosclerosis; ds.		
XX	Mus sp.		
XX	WO9818456-A1.		
XX	07-MAY-1998.		
XX	29-OCT-1997;	97WO-US19732.	
XX	30-OCT-1996;	96US-0742440.	
XX	(REGC) UNIV CALIFORNIA.		
XX	Connolly A, Coughlin SR, Ishihara H;		
XX	WPI; 1998-271905/24.		
XX	DNA encoding protease-activated receptor 3 - for detection of		
PT	specific agonists and antagonists, potentially useful for treating		
PT	e.g. thrombosis, atherosclerosis, inflammation etc.		
XX	Claim 3; Page 39; 74pp; English.		
XX	This sequence comprises genomic DNA for mouse protease-activated		
CC	receptor 3 (PAR3), a cell surface protein which is specifically		
CC	activated by thrombin or a thrombin agonist, thereby activating		
CC	signalling events such as phosphoinositide hydrolysis, calcium ion		
CC	efflux and platelet aggregation. A cDNA library was established		
CC	from rat platelets and amplified (primers given in AAV07376-77). A		
CC	product was identified that was expected to encode a new G-protein		
CC	coupled receptor, related to PAR1 or 2. This was used to isolate		
CC	human and murine DNA by a combination of PCR and hybridisation		
CC	techniques. Murine cDNA (see AAV07372), murine genomic DNA, human		
CC	cDNA (see AAV07374) and human genomic DNA (see AAV07375) sequences are		
CC	provided, as well as the murine (see AAW51405) and human (see AAW51406)		
CC	PAR3 proteins. Also claimed are vectors, host cells and an assay		
CC	device. Host cells are used to screen compounds for their ability		
CC	to act as agonists or antagonists of the effects of thrombin-PAR3		
CC	interaction. Agonists are used to treat wounds, thrombosis,		
CC	atherosclerosis, restenosis, inflammation and other thrombin		
CC	activated disorders. Antagonists (see AAW51415-21) are used to		
CC	control blood coagulation and thereby to treat heart attack and		
CC	stroke. They also mediate inflammatory and proliferative responses		
CC	to injury as occur in wound healing, atherosclerosis, restenosis,		
CC	pulmonary inflammation (ARDS) and glomerulosclerosis.		
XX	Sequence 1124 BP; 251 A; 314 C; 210 G; 326 T; 23 other;		
SQ	Pred. No.: 3.48e-79 . Length: 1124		
Alignment Scores:			
Pred. No.:	3.48e-79 . Length: 1124		

Score: 959.00 Matches: 185
Percent Similarity: 73.99% Conservative: 34
Best Local Similarity: 62.50% Mismatches: 53
Query Match: 44.46% Indels: 25
DB: 19 Gaps: 2

US-09-208-629F-6 (1-408) x AAV07373 (1-1124)

Qy 44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaPro 63
Db 36 AATGTTTCAGACAACCTCAGCAAGCCAACTTAATTAAGAGTTTAAATGGGGTCCC 95
Qy 64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
Db 96 CAAATACCTTTGAAGAATTCCNN----- 118
Qy 84 ThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
Db 119 -----NNNNTRCAAACTCTCCATGTAATGAATATGCTACC 151
Qy 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuVal 123
Db 152 ATGGGATACCTGAGAAGTTCCTTAAGTACCAAGTGATACCTGCCATCTACATCTGCTG 211
Qy 124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPheArgThrArg 143
Db 212 TTTGTGATGTGTGATGACGAGCAACATCGTGACCCCTGTGAAACTCTCTCAAGGACCAA 271
Qy 144 SerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaPheLeuPheCysVal 163
Db 272 TCCATCTGTGTCATCTTTTCACCACTGCGCATCGCGGATCTCTTTTCTGTGTC 331
Qy 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVa 183
Db 332 AGCTGCGGTTTAAGATCNCCTACCATCTCAATGGCAACAACTGGGTATTTGGCGAGT 391
Qy 183 IleCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSer---IleLeuLe 202
Db 392 CATGTGCGGATACCAACGGTCTGTTTTCACGGCAACATGCTACTGCGCTANNNTCTGAT 451
Qy 202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgG 222
Db 452 CCTCACTGTATGGGATCAACCGCTACCTGGCCACCGCTCACCTTTTCATACACAGAA 511
Qy 222 yLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
Db 512 GCTGCCCAACGAGCTTCTCCATGCTCATGTGTGGCATGGTGGGTCTATGTTTCTT 571
Qy 242 uTyrMetLeuProPhePheIle-LeuLysGlnGluTyrTyrLeuValGlnProAspIle 262
Db 572 ATACATGCTGCCCTTGTGTCATCCNNAAGCAGGAGTACCACCTCGTCCACTCCGAGATCA 631
Qy 262 hrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyrTrp 282
Db 632 CCACCTGCCAGATGTGTGACGCGTGGGANTCCCATCATCTCTCCGATTCTACTACT 691
Qy 282 heIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrA 302
Db 692 TCGTCTCTTAGCATCTTTGGGTCTCTCATCCGTTTGTGATCATCATCTTCTGTATACA 751
Qy 302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTyrTyrValLysAla 322
Db 752 CGACTCTCATCCCAAACTTAAATCAAAAGATCNGATATGGCTGGGTACATCAAGGCG 811
Qy 322 erLeuLeuLeuValIlePheThrIleCysPheAlaProSer 336
Db 812 TCTCTCTCATCTTGTGAATTTACCATCTGCTTCCCTCCACCACC 855

RESULT 10

AAQ84559

ID AAQ84559 standard; cDNA; 2732 BP.

XX

AC AAQ84559;

XX 22-AUG-1995 (first entry)
XX Murine C140 receptor cDNA.
XX G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX Mus musculus.
XX Key Location/Qualifiers
XX CDS 73..1272
XX /*tag= a
XX WO9503318-A.
XX PN 02-FEB-1995.
XX PD 26-JUL-1994; 94WO-US08536.
XX PF 26-JUL-1993; 93US-0097938.
XX PR (CORT-) COR THERAPEUTICS.
XX PI Scarborough RM, Sundelin J;
XX WPI; 1995-075182/10.
XX DR P-PSDB; AAR66922.
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX and antagonists and specific antibodies with therapeutic and
XX diagnostic applications.
XX Claim 1; Fig 10; 57pp; English.
XX A cDNA library from a mouse stomach was constructed in lambda gt10
XX and screened with a probe encompassing the C140 genomic clone (see
XX AAQ84557). A single phage clone was isolated and cut with EcoRI. The
XX insert was cloned in pBluescript and pSG5 and sequenced. The
XX complete nt sequence and deduced AA sequence is given in AAQ84559 &
XX AAR66922. 5' RACE resulted in the addition of only 27 bps to the 5'
XX end. The 5' end of the apparent coding region differs from the 5'
XX end of the ORF of genomic DNA; it is believed that the 5' end of
XX the cDNA sequence is correct.
XX SQ Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;
Alignment Scores:
Pred. No.: 3,98e-46 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 16 Gaps: 12
US-09-208-629F-6 (1-408) x AAQ84559 (1-2732)
Qy 22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db 85 AGCTGCGGTGGTCTGGGAGGTATCACCTTTCTGGCGGCTCTCTCTCCACGCCG 144
Qy 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db 145 ACCGAGAACCTTGACCGGAGCGCAACACAGTAAAGGAAGAGTCTTATTGGCAGATTA 204
Qy 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db 205 GAAACCCAGCCTCCA-----ATCACT 225
Qy 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db 226 GGGAAAGGGGTTCGGGTAGAACCCAGGCTTTTCCATCGATGAGTCTCTCGGTCC----- 279
Qy 98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIlePro 117

```
Db 280 -----ATCCTCACCGGAAGTGTACCAACGGTCTTTCTTCG 315
Qy 118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAlaValThrLeuTyrMet 137
Db 316 GTGCTGTACATATTGTTGTGATGTTGTCAGTAAATGCGCCCTCTGGATC 375
Qy 138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db 376 TTCCTTTCCGACGAGAGAAACACCCCGCGTGTATTACATGGCCCAACCTGGCGCTG 435
Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db 436 GCGACCTCTCTCTGTCTGTCTGTGTTCCCTCCAGATCTCTTACCACCTTACATGCAAC 495
Qy 177 AsnTyrValPheGlyGluValLeuLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 496 AACTGGGTCTACGGGGAGGCCCTGTGCAAGGTGCTCATTTGCTTTTCTATGTAACATG 555
Qy 197 TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db 556 TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGACAGGTACTGGGTGATCGTGAAC 615
Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
Db 616 CCCATG-----GGACACCCCGAGGAAG---GCAACATCGCGCTGGCGTCTCC 663
Qy 236 -----ValTyrAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
Db 664 TTGGCAATCTGGTCTGATTTCTGCTGCTCAGTCCCTTTGTATGTCATGAAGACGACC 723
Qy 254 TyrTyrLeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGluSer 273
Db 724 ATCTACATTCAGCATTTGAACATCACCATCTGTACAGATGTG----- 765
Qy 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
Db 766 ---CTGCTCTGAGGAGTATTGGTGGGGACATGTTCAATTACTTCTCTCACTGGCCATT 822
Qy 288 PheGlyPheLeuLeuProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db 823 GGAGTCTCTCTGTTCGGCGCCCTCTTACTGCTGCTACGCTGCTCATGATCAAGACG 862
Qy 308 LeuAsnAla-----TyrAspHisArgTrpLeuTyrTyrValLysAla 321
Db 883 CTCGCTCTCTGCTATGATGATACACTCAGAGAAAGGAGGCGGTATCCGACTC 942
Qy 322 SerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 341
Db 943 ATCATCACCGTGTGGCCATGTACTTCTCTGCTCTAGCAACCTTCTGCTCGTA 1002
Qy 342 IleHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db 1003 GTGCATTATTTCCTTAATCAAAACCCAGGACGACGACGCTACGCGCTCTACCTGTGTC 1062
Qy 362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1063 GCCCTCTGCTCTGACCTTCAACAGCTGTATAGACCCCTTGTCTATTACTTGTCTCA 1122
Qy 382 LysThr---ArgAsnHisSer 387
Db 1123 AAAGATTTTCAGGATCAGCC 1143
```

RESULT 11

AAT32038
ID AAT32038 standard; cDNA; 2732 BP.

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

XX AC AAT32038;

KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure; ss.

OS Mus sp.

XX Key Location/Qualifiers
FH CDS 73..1272

FT sig_peptide /tag= a

FT /tag= b

FT /note= "the signal sequence differs from a
signal sequence shown in a genomic DNA
segment encoding this receptor (see AAT32036),
the signal sequence given for this clone is
believed to be the correct sequence"

FT mat_peptide 166..1269

XX /tag= c

PN W09623225-A1.

XX 01-AUG-1996.

XX 25-JAN-1996; 96WO-US01179.

XX 25-JAN-1995; 95US-0390301.

XX (CORT-) COR THERAPEUTICS INC.

XX Scarborough RM, Sundelin J;

XX WPI; 1996-362813/36.

XX P-PSDB; AAW01954.

XX Vector for expression C140 cell surface receptor in host cell
PT useful to identify C140 agonist and antagonists, which are
PT antihypertensives and elevators of blood pressure, respectively
PS Example 4; Fig 10A-B; 60pp; English.

XX AAT32038 encodes the murine C140 receptor (C140R). The sequence may be
engineered so as to allow the recombinant expression of C140R in a
suitable host cell, i.e. by removing the native expression-control
sequence and replacing them with control sequences operable in the
host. Such a recombinant receptor can be expressed on the surface of
oocytes, this provides a good assay system for identifying
agonists/antagonists of C140R. The C140 receptor is a G-protein linked
receptor and a member of the "seven-pass" transmembrane receptor
superfamily (peptide chain of the receptor passes through the cell
membrane seven times, producing seven transmembrane regions within
the receptor molecule). The C140 receptor is involved in controlling
blood pressure. C140 antagonists (see AAW01942-W01951) are useful to
inhibit signalling from this receptor, resulting in an increase in
blood pressure and are therefore useful in pharmaceuticals for the
treatment of hypotension (low blood pressure). Conversely agonists
(see AAW01914-W01941) of C140 are useful in pharmaceuticals for the
treatment of hypertension (high blood pressure).

XX Sequence 2732 BP; 650 A; 676 C; 633 G; 773 T; 0 other;

Alignment Scores:

Pred. No.: 3,98e-46 Length: 2732
Score: 606.50 Matches: 137
Percent Similarity: 53.23% Conservative: 69
Best Local Similarity: 35.40% Mismatches: 126
Query Match: 28.12% Indels: 55
DB: 17 Gaps: 12

US-09-208-629F-6 (1-408) x AAT32038 (1-2732)

Qy 22 AlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGly 41
Db 85 AGCCTGGCTGGCTGCTGGGAGGTATCACCCCTTCTGCGCGCTCTCTCGACGCGG 144


```
Qy 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
Db 145 ACCGAGAACTTGCACCGGAGCAACAGTAAGGAAGAGTCTTATTGGCAGATTA 204
Qy 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
Db 205 GAAACCCAGCTCCA-----ATCACT 225
Qy 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
Db 226 GGGAAAGGGTTCGGTAGAACACAGGCTTTTCCATCGATGAGTCTCTGCGTCC-----279
Qy 98 HisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuPro 117
Db 280 -----ATCCTCACCGGAGAGCTGACACCGCTTCTTCTCCG 315
Qy 118 AlaIleTyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMet 137
Db 316 GTCGTCTACATTATTGTTGTGATTTGTTGTTGCCAGTAATGGCATGGCCCTCTGGATC 375
Qy 138 LeuPhePheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIle 156
Db 376 TTCCTTTCCGACGAAGAAGAACACCCCGCGTGATTTACATGGCCACCTGGCCTTG 435
Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsn 176
Db 436 GCGACCTCTCTCTGTCATCTGTTCCCTCCCTGAAGATCTCCTACCACTACATGGCAAC 495
Qy 177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMet 196
Db 496 AACTGGGTCTACGGGAGGCGCTGTGCAAGGTGCTCAFTGGCTTTTCTATGTAACATG 555
Qy 197 TyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
Db 556 TATGCTCCATCTCTTCATGACCTGCTGAGCGTGACAGGTACTGGGTGATCGTGAAC 615
Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu---235
Db 616 CCCATG-----GGACACCCCGAGGAAG--GCAACATCGCGCTGGCGTCTCC 663
Qy 236 -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLysGlnGlu 253
Db 664 TTGGCAATCTGGCTCTGATTTTCTGGTCACATCCCTTTGATGTCATGAAGCAGAC 723
Qy 254 TyrTyrLeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGluSer 273
Db 724 ATCTACATCCAGCATTTGAACATCACACCTGTCACGATGTG-----765
Qy 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
Db 766 ---CTGCTCAGGAGGTATTGGTGGGGGACATGTTCAATTACTTCTCTCACTGGCCATT 822
Qy 288 PheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThr 307
Db 823 GGAGTCTCTCTGTTCCCGCCCTCTTCTGTCATCTCGCTACGCTGCTCATGATCAAGC 882
Qy 308 LeuAsnAla-----TyrAspHisArgTrpLeuTyrTyrValLysAla 321
Db 883 CTCGCTCTCTGCTATGATGATGATACATCTCAGAGAGAAAGGAGGCTATCCGACTC 942
Qy 322 SerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAniIleLeuIle 341
Db 943 ATCATCACCGTGGCGCATGTACTTCTCTGCTTCTCCTAGCAACCTTCTGCTGTA 1002
Qy 342 IleHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIle 361
Db 1003 GTGCATTATTTCCTAATCAAACCCAGGAGCAGCCACGCTCTACGCGCTCTACCTTGT 1062
Qy 362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1063 GCCCTCTGCTGTGACCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCA 1122
Qy 382 LysThr---ArgAsnHisSer 387
```

```
Db 1123 AAAGATTTCAGGATCAGGCC 1143
RESULT 12
AAQ84557
ID AAQ84557 standard; DNA; 1477 BP.
XX
AC AAQ84557;
XX
DT 22-AUG-1995 (first entry)
XX
DE Murine C140 receptor genomic DNA.
XX
KW G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
OS Mus musculus.
XX
PH Key Location/Qualifiers
FT CDS 232..1188
/*tag= a
XX
XX WO9503318-A.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US08536.
XX
XX 26-JUL-1993; 93US-0097938.
XX
XX (CORT-) COR THERAPEUTICS.
XX
XX Scarborough RM, Sundelin J;
XX
XX WPI: 1995-075182/10.
XX
XX P-PSDB; AAR66920.
XX
XX New DNA encoding recombinant C140 receptor - and novel agonists
XX PT and antagonists and specific antibodies with therapeutic and
XX PT diagnostic applications.
XX
XX Disclosure; Fig 1; 57pp; English.
XX
XX A mouse cosmid genomic library (obtd. from Dr R.A. Wetsel, Washington
XX Univ. School of Medicine, St Louis, Missouri) was screened with two
XX 32P-labeled oligos corresp. to bp 190-249 and 742-801 of the bovine
XX substance K receptor cDNA. In one of the clones isolated (C140) the
XX hybridising region was localised to a 3.7 kb PstI fragment. This
XX fragment was subcloned into pBluescript vector. The hybridising and
XX CC adjacent regions were sequenced. The nt sequence and the deduced AA
XX CC sequence are given in AAQ84557 & AAR69920 respectively. The published
XX CC DNA sequence appears to have been printed with the last two lines
XX CC in the wrong order. This has been corrected in this GeneSeq entry.
XX
SQ Sequence 1477 BP; 341 A; 414 C; 320 G; 402 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 6,44e-46 Length: 1477
Score: 600.50 Matches: 139
Percent Similarity: 52.25% Conservative: 70
Best Local Similarity: 34.75% Mismatches: 118
Query Match: 27.84% Indels: 74
DB: 16 Gaps: 13
```

US-09-208-629F-6 (1-408) x AAQ84557 (1-1477)

```
Qy 18 IleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuProThr--- 36
Db 221 GTTCATCTCAATGTTCCATTTAAACACACAGCAGCCTTACTGTTGACCATTTATCTCAG 280
Qy 37 -----PhcCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
Db 281 TAATGATTCTGCTCGCTTTCTTTGTACAGGAGCGCAACACAGTAAGGAAGTCTTTA 340
```

```
QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
Db 341 TTGGCAGATTAGAAACCCAGCTCCCAATCACTACGGAAGGGTTCGGTA-GAACAGGC 399
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db 400 TTTTCCATCATGATGATTC-----417
QY 85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db 418 -----TTCGCTCC-----426
QY 105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db 427 ---ATCTCACCGGAAGCTGACCAAGCTTCTTCCGTCGCTCTACATTATGTGTT 483
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db 484 GTGATTGGTTTGCCAGTAATGGCATGGCCCTCTGGATCTTCTTTCCGACGAAGAAG 543
QY 145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db 544 AAACACCCCGCCGCTGATTTACATGGCCAACTGGCCCTGGCCGACCTCTCTCTGTCATC 603
QY 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db 604 TGGTCCCTCGAAGATCTCTACCACTACATGGCAACACTGGGTCTACGGGGAGGCC 663
QY 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db 664 CTGTGCAAGGTGCTCAATGGCTTTTCTATGGTAACATGATTGCTCCATCTCTTCATG 723
QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db 724 ACCTGCTCAGCGCGAGAGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTyrAlaThrVal 240
Db 775 CCCAGGAAGAG---GCAACATCGCGCTGGGTCTCTTGGCAATCTGGCTCTCTGATT 831
QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db 832 TTTCTGCTCACCATCCCTTTGTATGTATGAAGCAGACCATCTACATTCAGCATGGAAC 891
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db 892 ATCACCACCTGTACAGATG-----CTGCCCTGAGGAGGTATTG 930
QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db 931 GTGGGGGACATGTTCAATTACTTCTCTACTGGCCATGGAGTCTCTGTTCGGGCC 990
QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----310
Db 991 CTCCTTACTGTCATGCTGCTCATGATCAAGACGCTCGCTCTCTCTGCTATGGAT 1050
QY 311 -----TyrAspHisArgTyrPheTyrValLysAlaSerLeuLeuValIle 328
Db 1051 GAACACTCAGAGACAAAGGACAGAGGCTATCGACTCATCATCGCTGGCCATG 1110
QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleHisHisAlaAsnTyrTyr 348
Db 1111 TACTTCTATGCTTTCTGCTTAGCAACCTTCTGCTGATGGCATTTATTCCTAATCAA 1170
QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db 1171 ACCCAGAGGACAGCCAGCTACGCTCTACCTTGTGCGCTCTGCTGTCAGCCCTC 1230
QY 369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr----ArgAsnHisSer 387
Db 1231 AACAGCTGCATAGACCCCTTTGTCTATTACTTTGTCTCAAAAGATTTCAGGGATCAGCC 1290
, RESULT 13
```

```
AAT32036
ID AAT32036 standard; DNA; 1477 BP.
AC AAT32036;
XX 01-APR-1997 (first entry)
DT Murine C140 receptor-encoding DNA.
DE C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ds.
XX Mus sp.
OS Location/Qualifiers
FH CDS 232..1419
FT /*tag= a
FT sig_peptide 232..312
FT /*tag= b
FT /*note= "putative signal sequence, differs from
FT signal sequence shown in a cDNA clone of
FT this receptor (see AAT32038), the signal
FT sequence given for the cDNA clone is
FT believed to be the correct sequence"
FT mat_peptide 313..1416
FT /*tag= c
XX WO9623225-A1.
XX 01-AUG-1996.
XX 25-JAN-1996; 96WO-US01179.
XX 25-JAN-1995; 95US-0390301.
XX (CORT-) COR THERAPEUTICS INC.
XX Scarborough RM, Sundellin J;
XX WPI; 1996-362813/36.
XX P-PSDB; AAW01952.
XX Vector for expression C140 cell surface receptor in host cell
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX Example 1; Fig 1A-B; 60pp; English.
XX AAT32036 encodes the murine C140 receptor (C140R), including a tentative
XX signal sequence. The sequence may be engineered so as to allow the
XX recombinant expression of C140R in a suitable host cell, i.e. by
XX removing the native expression-control sequence and replacing them with
XX control sequences operable in the host. Such a recombinant receptor
XX can be expressed on the surface of oocytes, this provides a good assay
XX system for identifying agonists/antagonists of C140R. The C140
XX receptor is a G-protein linked receptor and a member of the "seven-
XX pass" transmembrane receptor superfamily (peptide chain of the
XX receptor passes through the cell membrane seven times, producing
XX seven transmembrane regions within the receptor molecule). The C140
XX receptor is involved in controlling blood pressure. C140 antagonists
XX (see AAW01942-W01951) are useful to inhibit signalling from this
XX receptor, resulting in an increase in blood pressure and are therefore
XX useful in pharmaceuticals for the treatment of hypotension (low blood
XX pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
XX in pharmaceuticals for the treatment of hypertension (high blood
XX pressure).
XX Sequence 1477 BP; 341 A; 415 C; 319 G; 402 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 6.44e-46 Length: 1477
XX Score: 600.50 Matches: 139
XX Percent Similarity: 52.13% Conservative: 69
```

Best Local Similarity: 34.84% Mismatches: 120
Query Match: 27.84% Indels: 72
DB: 17 Gaps: 12

US-09-208-629F-6 (1-408) x AAT32036 (1-1477)

QY 18 IleLysMetLysAlaLeuLeuPheAlaAlaGlyLeuLeuLeuLeuLeuProThr--- 36
DB 221 GTTCATCTGAAATGTTCCATTATAAACACAGCAGCGCTTACTGTTGGACCATTTATCTCAG 280
QY 37 -----PheCysGlnSerClyMetGluAsnAspThrAsnAsnLeuAla----- 50
DB 281 TAATGATTCTGCTCCGCTTTCTTTGTACAGCAGCAACACAGTAAGGAAGAGAGTCTTA 340
QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
DB 341 TTGGCAGATTAGAAACCCAGCTCCATCACTGGGAAGGGGTTCCGGTA-GAACCAGGC 399
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
DB 400 TTTTCCATCGATCAGTTC----- 417
QY 85 ValLysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
DB 418 -----TCTGCGTCC----- 426
QY 105 GlyTyLeuThrSerSerLeuLeuLeuLeuProAlaIleTyLeuLeuValPhe 124
DB 427 ---ATCCTCACCGGAGCTGACACGGCTTTCTTCGGTGGTCTACATTATTGTGTTT 483
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
DB 484 GTGATTGGTTGCCAGTAATGGCATGGCCCTCTGGATCTTCTTTCCGAACGAAGAAG 543
QY 145 IleCysThrThrValPheTyThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
DB 544 AAACACCCCGCGGTGATTACATGGCCAAACCTGGCGTTGGCGGACCTCTCTGTCATC 603
QY 164 ThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
DB 604 TGGTCCCTCGAGATCTCTACACCATCATGCGCAACACTGGGTCTACGGGAGGCC 663
QY 184 LeuCysArgAlaThrThrValIlePheTyGlyAsnMetTyCysSerIleLeuLeuLeu 203
DB 664 CTGTGCAAGGTGCTCATTTGCTTTTCTATGTTAATCATGTTATGCTCCATCTCTTCATG 723
QY 204 AlaCysIleSerIleAsnArgTyLeuAlaIleValHisProPheThrTyArgGlyLeu 223
DB 724 ACCTGCTCAGGCTGAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
QY 224 Pro-----LysHisThrTyAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
DB 775 CCAGGAAGAGGCAACATCCCGTGGCTCTCTCTGGCAATCTGGCTCTCTGATTTT 834
QY 242 LeuTyMetLeuProPheIleLeuLysGlnGluTyTyLeuValGlnProAspIle 261
DB 835 CTGGTCACCATCTTTGATGTCATGAAGCAGACCATCTACATTCCAGCATTCGAATC 894
QY 262 ThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyThr--- 280
DB 895 ACCACCTGTCAGATG-----CTGCGTGGAGGAGGTATTGGTG 933
QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuLeuProPheVal 295
DB 934 GGGGACATGTTCAATTACTTCTCTACATGGCCATTTGGAGTCTTCTCTCCCGCCCTC 993
QY 296 LeuIleIleTyCysTyThrAlaAlaIleIleArgThrLeuAsnAla----- 310
DB 994 CTTACTGCATCTCGCTACGTCATGATCAGACGCTCCGCTCTCTGCTATGATGAA 1053
QY 311 ---TyrAspHisArgTrpLeuTrpTyValLysAlaSerLeuLeuLeuValIlePhe 329
DB 1054 CACTCAGAGAACAAAGGAGGAGGCTATCCGACTCATCATCACCCTGCTGGCCATGATC 1113

QY 330 ThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyTyTyTyr 349
DB 1114 TTCACTGCTTTTCGTCCTAGCAACCTTCTGCTGCTAGTCATTATTTCTTAATCAAAACC 1173
QY 350 AsnAsnThrAspGlyLeuTyTyPheIleTyLeuIleAlaLeuCysLeuGlySerLeuAsn 369
DB 1174 CAGAGCGAGAGCCAGCTACGCGCTCTACCTTCTGCGCTCTGCTGCTGCTCAAC 1233
QY 370 SerCysLeuAspProPheLeuTyTyPheLeuMetSerLysThr---ArgAsnHisSer 387
DB 1234 ACCTCATAGACCCCTTGTCTATTACTTGTCTCAAGATTTCAGGGATCAGCC 1290

RESULT 14
AAC60319
ID AAC60319 standard; DNA; 1451 BP.
XX AAC60319;
AC AAC60319;
DT 19-FEB-2001 (first entry)
XX Human PAR-2 DNA.
DE
XX PAR-2; protease activated receptor-2; ECL-2; inflammatory disease;
KW asthma; chronic obstructive pulmonary; arthritis;
KW inflammatory bowel; psoriasis; eczema; multiple sclerosis; ds.
XX Homo sapiens.
OS
XX WO200063371-A1.
PN
XX 26-OCT-2000.
PD
XX 17-APR-2000; 2000WO-GB01455.
PF
XX 15-APR-1999; 99GB-0008513.
PR
XX (UYSO-) UNIV SOUTHAMPTON.
PA
XX Wallis AF, Palmer K, Compton SJ, Cairns JA, Gough AC;
PI WPI; 2000-679599/66.
XX
DR Protease activated receptor 2 variants useful for treating inflammatory
XX diseases such as asthma, arthritis and psoriasis, and as hypotensives,
XX PT has reduced sensitivity to trypsin -
XX
XX Disclosure; Page 54-55; 59pp; English.
XX
CC The present invention relates to a variant protease activated
XX receptor 2 (PAR-2). The invention is useful for identifying an
XX individual having a polymorphism in the ECL-2 region of one or
XX both PAR-2 gene alleles. The invention may be used to develop treatments
XX for inflammatory diseases such as asthma, chronic obstructive
XX pulmonary diseases, arthritis, inflammatory bowel diseases, psoriasis
XX and eczema, multiple sclerosis and to raise blood pressure.
XX
SQ Sequence 1451 BP; 310 A; 389 C; 346 G; 406 T; 0 other;

Alignment Scores:
Pred. No.: 1.24e-44 Length: 1451
Score: 586.50 Matches: 141
Percent Similarity: 50.61% Conservative: 67
Best Local Similarity: 34.31% Mismatches: 146
Query Match: 27.19% Indels: 57
DB: 21 Gaps: 14

US-09-208-629F-6 (1-408) x AAC60319 (1-1451)

QY 8 Ser***ArgLeuArgAspGlyThrClnValIleLysMetLysAlaLeuIlePheAlaAla 27
DB 139 TCCAGGAGGATCGGAGCCCGCGCGGTGGTGTCTGCTGGCGGCCCATCTCTCTAGCA 198

1054 CACTCAGAGAACAAAGGCAGAGGGCTATCCGACTCATCATCCGTCGTGGCCATGTAC 1113


```
Db 260 -----GTCCTCAGTGGAAACTGACCACGTGCTCCCTTCCAATGTC 301
Qy 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 302 TACAACATTTGTTGTGGTGGTTTGCAGTAACGGCATGGCCCTGTGGGTCTTCTT 361
Qy 140 PheArgThrArgSerIleCysThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 362 TTCGAACATAAGAAAGAACCCCTGCTGTGATTTACATGCGCAATCTGGCCTGGCTGAC 421
Qy 159 PheLeuPheCysValThrLeuProPheIleAlaTyrHisLeuAsnGlyAsnAspTyr 178
Db 422 CTCCTCTCTGTCATCTGGTCCCTTGAAGATTCCCTATCACATACATGCAACAACCTGG 481
Qy 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db 482 ATTATGGGAAGCTCTTTGTAATGCTTATTGGCTTTTCTATGGCAACATGACTGT 541
Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 542 TCCATTCTCTTCATGACCTGCTCAGTGTGCAGAGGTATTGGGTGTCATCGTGAACCCCATG 601
Qy 219 ThrTyrArgGlyLeuProIleHisThrTyrAlaLeuValThrCysGlyLeuValTyrAla 238
Db 602 GGGCAC---TCCAGGAAGAGGCAACATTGCCATTCGCCATCTCCCTGGCAATATGGCTG 658
Qy 239 ThrValPheLeuTyrMetLeuProPheIleLeuLeuGlnGluTyrTyrLeuValGln 258
Db 659 CTGATTCGTGGTCACCATCCCTTTGTATGTGCTGAGGAGACCATCTTCATTCCTGCC 718
Qy 259 ProAspIleThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 719 CTGAACATCAGACCTGTCATGATGTT-----TTGCCCTGAGCAG 757
Qy 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeu 292
Db 758 CTCCTGGTGGAGACATGTTCAATTACTCTCTCTGCGCAATGGGGTCTTCTGTTTC 817
Qy 293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
Db 818 CCAGCCTCTCTCACAGCCTCGCTATGCTGTGATGATCAGATGCTGCATCTCTGCTGCC 877
Qy 311 -----TyrAspHisArgTyrLeuTyrValIleValSerLeuLeuLeu 326
Db 878 ATGATGAAAACTCAGAGAAGAAAGGAGGCGCATCAAACTCATTTGCTCACTGCTGCTG 937
Qy 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn 346
Db 938 GCCATGTACCTGATCTCTTCACTCTCCTAGTAACCTTCTGCTGGTGGCAT----- 988
Qy 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 989 TATTTCTGATTAGACCGAGCGGCGCAGCGCATGCTATGCCCTGTCATTTAGCCCTC 1048
Qy 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 1049 TGCCTCTCTACCCCTTAACAGCTGATGATGACCCCTTTGCTATTACTTTGTTTTCACATGAT 1108
Qy 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1109 TTCAGGGATGATGCAAGAACGCTCTCTTGGCGAAGTGTGCGCACTGTAAAGCAG 1165
RESULT 16
AAT32037
ID AAT32037 standard; DNA; 1255 BP.
XX AAT32037;
XX 01-APR-1997 (first entry)
XX Human C140 receptor-encoding DNA.
XX
```

```
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
XX antagonist; hypertension; hypotension; blood pressure; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 56..1197
FT /*tag= a
FT sig_peptide 56..136
FT /*tag= b
FT /*note= "putative signal sequence, differs from
FT signal sequence shown in a cDNA clone of
FT this receptor (see AAT32039 - this clone also
FT encodes a mature receptor having four amino
FT acid changes, when compared to this sequence)"
FT mat_peptide 137..1194
FT /*tag= c
XX WO9623225-A1.
XX 01-AUG-1996.
XX 25-JAN-1996; 96WO-US011179.
XX 25-JAN-1995; 95US-0390301.
XX (CORT-) COR THERAPEUTICS INC.
XX Scarborough RM, Sundelin J;
XX WPI; 1996-362813/36.
XX P-PSDB; AAW01953.
XX Vector for expression C140 cell surface receptor in host cell -
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX Example 2; Fig 2A-B; 60pp; English.
XX AAT32037 encodes the human C140 receptor (C140R), including a tentative
XX signal sequence. The sequence may be engineered so as to allow the
XX recombinant expression of C140R in a suitable host cell, i.e. by
XX removing the native expression-control sequence and replacing them with
XX control sequences operable in the host. Such a recombinant receptor
XX can be expressed on the surface of oocytes, this provides a good assay
XX system for identifying agonists/antagonists of C140R. The C140
XX receptor is a G-protein linked receptor and a member of the "seven-
XX pass" transmembrane receptor superfamily (peptide chain of the
XX receptor passes through the cell membrane seven times, producing
XX seven transmembrane regions within the receptor molecule). The C140
XX receptor is involved in controlling blood pressure. C140 antagonists
XX (see AAW01942-W01951) are useful to inhibit signalling from this
XX receptor, resulting in an increase in blood pressure and are therefore
XX useful in pharmaceuticals for the treatment of hypotension (low blood
XX pressure). Conversely agonists (see AAW01914-W01941) of C140 are useful
XX in pharmaceuticals for the treatment of hypertension (high blood
XX pressure).
XX SQ Sequence 1255 BP; 294 A; 320 C; 260 G; 381 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1,27e-44 Length: 1255
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 17 Gaps: 10
```

US-09-208-629F-6' (1-408) x AAT32037 (1-1255)

```
Qy 80 GlyAlaThrIleThrValIleLysCysProGluGluSerHisLeuHisVal 99
Db 212 GGAGTTACAGTTGAACACAGTCTTTTCTGTGGATGAGTTTCTGTCATCT----- 259
```



```

Qy 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrIleuAlaIleValHisProPhe 218
Db 498 TCCATTCTTCATGACCTGCTCAGTGTGCAGAGGTATTGGGTCTATCGTGAACCCCATG 557
Qy 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrAla 238
Db 558 GGGCAC---TCCAGGAAGAGGCAACATTCGCCATTCGCCATTCCTCCCTGGCAATATGCTG 614
Qy 239 ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 615 CTGATTCTGGTGGTCCACATCCCTTTGTATGTCTGTGAGGAGACCATCTTCATCTCGGCC 674
Qy 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 675 CTGAACATCAGACCTGCTCATGTGTT-----TTGCCTGAGCAG 713
Qy 279 LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db 714 CTCCTGGTGGGAGACATGTTCAATTACTTCTCTCTGCGCATTTGGGTCTTTCTGTTT 773
Qy 293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
Db 774 CCAGCCTTCTTCACAGCCTCTGCTATGTCTGTATGATCAGATGCTGCGCATCTCTGCC 833
Qy 311 -----TyrAspHisArgTyrLeuTyrTyrValLysAlaSerLeuLeuIleLeu 326
Db 834 ATGGATCAAAACTCAGAGAAGAAAGAGGAGGCCATCAAACTCATTTGTCACGTCTCTG 893
Qy 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHisAlaAsn 346
Db 894 GCCATGTACCTGATCTGCTTCACTCTCTAGTAACCTTCTGTGTGTGTCAT----- 944
Qy 347 TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db 945 TATTTTCTGATTAAAGCCAGCCAGCCAGCCATGCTATGCCCTGTACATTTGATGCCCTC 1004
Qy 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 1005 TGCCTCTCTACCTTAAACAGCTGATCGATCGACCCCTTTGTCTATTTACTTTTTCACATGAT 1064
Qy 384 ---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1065 TTCAGGATCATGCAAAAGAGCTCTCTCTTTCGCGAAGTGTCCGCACTGTAAAGCAG 1121

RESULT 18
AAQ84560
ID AAQ84560 standard; cDNA; 1414 BP.
XX
AC AAQ84560;
XX
XX 22-AUG-1995 (first entry)
XX
XX Human C140 receptor cDNA.
XX
XX G-protein-coupled receptor; G-protein; C140 receptor; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 50..1243
XX /*tag= a
XX
XX WO9503318-A.
XX
XX 02-FEB-1995.
XX
XX 26-JUL-1994; 94WO-US08536.
XX
XX 26-JUL-1993; 93US-0097938.
XX
XX (CORT-) COR THERAPEUTICS.
XX
XX

```

PI Scarborough RM, Sundelin J;

XX WPI; 1995-075182/10.

DR P-PSDB; AAR66923.

XX

PT New DNA encoding recombinant C140 receptor - and novel agonists
PT and antagonists and specific antibodies with therapeutic and
PT diagnostic applications.

XX

PS Claim 1; Fig 11; 57pp; English.

XX

CC A human intestinal tumour cDNA library was subjected to PCR using
CC primers designed from the genomic clone (see AAQ84558) and the
CC amplified fragment was cloned in pSG5 and sequenced. There are
CC four AA differences between the cDNA encoded sequence and that
CC encoded by the genomic DNA. The genomic DNA sequence and deduced
CC AA sequence are given in AAQ84560 & AAR66923.

XX

SQ Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;

Alignment Scores:

Pred. No.: 2,35e-43 Length: 1414
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 16 Gaps: 14

US-09-208-629F-6 (1-408) x AAQ84560 (1-1414)

```

Qy 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db 41 TCCAGGAGGATGGGAGCCCGCGGCTGCTGCGGGCGCCGATCTCTCTAGCA 100
Qy 28 AlaGlyLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db 101 GCCTCTCTC-----TCTGTC---AGTGGCACCATCCCAAGGAACCAAT 139
Qy 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProAsnSerPhe 67
Db 140 AGATCTCTTAAGGAAGAGCCCTTATTGTAGGTGTATGGCACA----- 184
Qy 68 GluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrValIle 87
Db 185 -----TCCACAGTCACTGGA---AAAGGAGTTACAGTTGAACAGACTTTT 226
Qy 88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db 227 TCTGTGGATGAGTTTCTGCATCT-----GTCCCTC 256
Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
Db 257 GCTGGAACCTGACCACTGTCTTCTTCCCAATGTCTACCAATTTGTTGCGGTGGT 316
Qy 128 ValProAlaAsnAlaValThrLeuTyrMetLeuPheArgThrArgSerIleCysThr 147
Db 317 TTGCCAAGTAACGGCATGCCCTATGGTCTTTCTTTCGGAACTAAGAAGAAGCACCT 376
Qy 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 377 GCTGTGATTACATGGCCCAATCTGGCTTTGGTGTGACCTCTCTCTGTCTGTGTTCCC 436
Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValLeuCysArg 186
Db 437 TTGAAGATTGCCTTATCACATACATGCAACCACTGGATTTATGGGAAGCTCTTTGTAAT 496
Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 497 GTGCTTATTGGCTTTTCTTCTATCGCAACATGACTGTTCATCTCTTCATGACCTGCCTC 556
Qy 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db 557 AGTGTGCAGAGGTATTGGGTTCATCGTGAACCCCATCGGCGCAC---TCCAGGAAGAGGCA 613

```

```

Qy 227 ThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeuTyrMetLeuPro 246
Db 614 AACATTGCCATTGGCATCTCCCTGGCAATATGGCTGCTGACTCTGTGGTCACCATCCCT 673
Qy 247 PhePheLeuLeuGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db 674 TTGTATGCTGTAAGCAGACCATCTTCATTCTCTGCGCCCTGAACATCAGACCTGTCTGAT 733
Qy 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr----- 280
Db 734 GTT-----TTGCCCTGAGCAGCTCTTGTGGGAGACATGTTCAAT 772
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 773 TACTTCTCTCTGCGGCATGTTGGTCTTCTGTCCAGCCCTCTCTCAGAGCTCTGCC 832
Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 833 TATGTGCTGATGATCAGAACTGTCGATCTTCTGCCATGGATGAAAACTCAGAGAGAAA 892
Qy 315 TrpLeuTrpTyrValIysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
Db 893 AGGAAGAGGCCCATCAAACTCATTTGTCACCTGCTCTGGGCATGTACCTGATCTCTTCACT 952
Qy 335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 953 CCTAGTAACCTCTGCTGTGGTGCA-----TATTTCTGTATTAGACCCAGGCC 1003
Qy 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1004 CAGAGCCATGCTATGCTGCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1063
Qy 372 LeuAspProPheLeuTyrPheLeuMetSerIysThr----ArgAsnHisSer---ThrAla 389
Db 1064 ATCAGCCCTTGTCTATTCTTGTTCATGATTTCAGGATCATGCAAGATGCAAGAACGCT 1123
Qy 390 TyrLeuThrIysLys*****AsnAspLeuArgGlnGlnGlnPro 404
Db 1124 CTCCTTGGCGAGTGTCCGACCTGTAAAGCAGATGCAAGTACCC 1168
RESULT 19
AAT32039
ID AAT32039 standard; cDNA; 1414 BP.
XX
AC AAT32039;
XX
DT 02-APR-1997 (first entry)
XX
DE Human C140 receptor cDNA clone.
XX
KW C140 receptor; G-protein linked; coupled; seven pass; agonist;
KW antagonist; hypertension; hypotension; blood pressure; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 50..1243
FT FT /*tag= a
FT FT 50..130
FT FT /*tag= b
FT FT /note= "the signal sequence, differs from the
FT FT signal sequence shown in a genomic clone of
FT FT this receptor (see AAT32037). The cDNA clone
FT FT also encodes a mature receptor having four
FT FT amino acid changes, when compared to the
FT FT genomic clone"
FT FT mat_peptide 131..1240
FT FT /*tag= c
XX
FN W09623225-A1.
XX
PD 01-AUG-1996.

```

```

XX 25-JAN-1996; 96WO-US01179.
XX
XX 25-JAN-1995; 95US-0390301.
XX
XX (CORT-) COR THERAPEUTICS INC.
XX
XX Scarbrough RM, Sundelin J;
XX
XX WPI; 1996-362813/36.
XX
XX P-PSDB; AAW01955.
XX
XX Vector for expression C140 cell surface receptor in host cell -
XX useful to identify C140 agonist and antagonists, which are
XX antihypertensives and elevators of blood pressure, respectively
XX
XX Example 5; Fig 11A-B; 60pp; English.
XX
XX AAT32039 encodes the human C140 receptor (C140R). The sequence may be
XX engineered so as to allow the recombinant expression of C140R in a
XX suitable host cell, i.e. by removing the native expression-control
XX sequence and replacing them with control sequences operable in the
XX host. Such a recombinant receptor can be expressed on the surface of
XX oocytes, this provides a good assay system for identifying
XX agonists/antagonists of C140R. The C140 receptor is a G-protein linked
XX receptor and a member of the "seven-pass" transmembrane receptor
XX superfamily (peptide chain of the receptor passes through the cell
XX membrane seven times, producing seven transmembrane regions within the
XX receptor molecule). The C140 receptor is involved in controlling blood
XX pressure. C140 antagonists (see AAW01942-W01951) are useful to inhibit
XX signalling from this receptor, resulting in an increase in blood
XX pressure and are therefore useful in pharmaceuticals for the treatment
XX of hypotension (low blood pressure). Conversely agonists (see AAW01914-
XX AAW01941) of C140 are useful in pharmaceuticals for the treatment of
XX hypertension (high blood pressure).
XX
XX Sequence 1414 BP; 335 A; 361 C; 309 G; 409 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,35e-43 Length: 1414
Score: 572.50 Matches: 140
Percent Similarity: 49.64% Conservative: 66
Best Local Similarity: 33.73% Mismatches: 152
Query Match: 26.54% Indels: 57
DB: 17 Gaps: 14
US-09-208-629f-6 (1-408) x AAT32039 (1-1414)
Qy 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMethLysAlaLeuIlePheAlaAla 27
Db 41 TCCAGGAGGATGCGGAGCCCGAGCGGCTGCTGGGGCGCCCATCTCTGTAGCA 100
Qy 28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db 101 GCCTCTCTC-----TCTGTC-----AGTGGCACCATTCCAAAGAACCAAT 139
Qy 48 AsnLeuAlaIysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db 140 AGATCTCTAAAGGAAGAACCTTATGTGTAAGTTGATGGCACA----- 184
Qy 68 GluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrValIysIle 87
Db 185 -----TCCCACGTCACCTGGA---AAAGGAGTTACAGTTGAACAGATCTTT 226
Qy 88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db 227 TCTGTGGATGAGTTTCTGTCATCT-----GTCCCTC 256
Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
Db 257 GCTGGAAACTGACCACTGTCTCTCTCCCAATTTGTCTACCAATTTGTTGGGGGGT 316
Qy 128 ValProAlaAsnAlaValThrLeuThrLeuThrPheArgThrArgSerIleCysThr 147

```


Db 317 TTGCAAGTAAAGGCGCATGGCTCTTTTCCGAATAAGAAAGACACCCCT 376
 Qy 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
 Db 377 GCTGTGATTACATGGCCATCTGGCTTGGCTGACCTCTCTCTGTCATCTGGTTCCTCC 436
 Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
 Db 437 TTGAAGATTGCTTATCATACATACATGCGCAACAACTGGATTATGGGAAGCTCTTTGTAAT 496
 Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
 Db 497 GTGCTTATGGCTTTTCTATCGCAACATGACTGTTCTTCATTCATGACCTGCCTC 556
 Qy 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
 Db 557 AGTGTGAGAGTATTGGGTCTATCGTGAACCCATGGGGCAC---TCCAGGAAGAGGCA 613
 Qy 227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
 Db 614 AACATTGCCATTGCCATCTCCCTGGCAATATGGCTGCTGACTCTGCTGGTCACCATCCCT 673
 Qy 247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
 Db 674 TTGTATGTCGTGAAGCAGACCATCTTCATTCTCCCTGACATCATCAGACCTGTGATGAT 733
 Qy 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr----- 280
 Db 734 GTT-----TTGCTGAGCAGCTCTGTGGGAGACATGTTCAAT 772
 Qy 281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
 Db 773 TACTTCTCTCTGGCCATCTGGGTCTTTCTGTTCACGCTTCCTTCACAGCCTCTGCC 832
 Qy 301 TyrAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
 Db 833 TATGTGCTGATGATCAGAACTGCTGCGATCTTCTGCCATGGATGAAACTCAGAGAAGAAA 892
 Qy 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
 Db 893 AGAAGAGGGCCATCAACTCATGTGCTGCTGGCGATGACCTGATCTGCTTCACT 952
 Qy 335 ProSerAsnIleIleLeuIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
 Db 953 CCTAGTAACCTCTGCTTGTGTGCAT-----TATTTCTGATTAGAGCCAGGCG 1003
 Qy 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
 Db 1004 CAGAGCCATGCTATGCGCTGTACATGTAGCCCTCTGCTCTCTACCTTAACAGCTGC 1063
 Qy 372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
 Db 1064 ATGACCCCTTGTCTATTACTTTGTTTTCATGATTTTCAGGGATCATCAAAAGAGCCT 1123
 Qy 390 TyrLeuThrLys*****AsnAspLeuArgGlnGlyGlnPro 404
 Db 1124 CTCCTTTGCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACC 1168

RESULT 20

AAx90979

ID AAX90979 standard; cDNA; 1361 BP.

XX AC AAX90979;

XX XT 25-JAN-2000 (first entry)

XX DE Mouse protease-activated receptor 4 (PAR4) cDNA.

KW Mouse protease-activated receptor 4 cDNA; PAR4; thrombin receptor;
 KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
 KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
 KW wound; blood coagulation; heart attack; stroke; inflammatory response;

KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
 KW pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.

Mus musculus.

Key Location/Qualifiers
 CDS 1..1191
 FT /*tag= a
 FT /product= "PAR4"

PN WO9943809-A2.

XX 02-SEP-1999.

XX 11-FEB-1999; 99WO-US02983.

XX 27-FEB-1998; 98US-0032397.

XX (REGC) UNIV CALIFORNIA.

XX Coughlin SR, Kahn M;

XX WPI; 1999-619953/53.

XX P-PSDB; AAY15081.

XX DNA molecules encoding protease-activated receptor 4, useful in

XX compound assays for thrombin agonist and antagonist activity -

XX Claim 3; Fig 1; 69pp; English.

XX The present sequence is a cDNA encoding protease-activated receptor 4

XX (PAR4) and derived from a mouse 14-15 days embryo library. PAR4 is a G

XX protein-coupled thrombin receptor expressed on cell surface. It is

XX activated by thrombin and mediates signalling events e.g.

XX phosphoinositide hydrolysis, calcium efflux and platelet aggregation.

XX The receptor is highly expressed in spleen cells and likely to be

XX involved in thrombin-mediated activation of platelets and other

XX haematopoietic cells. It is used for screening novel thrombin agonists

XX or antagonists. The agonists are used as therapeutics to treat wounds,

XX promote clotting and as reagents to activate platelets in diagnostic

XX tests. Antagonists are used to control blood coagulation, treat heart

XX attacks and strokes, and block inflammatory and proliferative responses

XX that occur in normal wound healing and variety of diseases including

XX atherosclerosis, restenosis, pulmonary inflammation (ARDS) and

XX glomerulosclerosis.

XX SQ Sequence 1361 BP; 234 A; 416 C; 394 G; 317 T; 0 other;

Alignment Scores:

Pred. No.: 1.37e-39 Length: 1361

Score: 531.50 Matches: 112

Percent Similarity: 53.82% Conservative: 57

Best Local Similarity: 35.67% Mismatches: 128

Query Match: 24.64% Indels: 17

DB: 20 Gaps: 2

US-09-208-629F-6 (1-408) x AAX90979 (1-1361)

Qy 84 ThrValLysIleLysCysProGluGluSer----- 93

Db 130 ACAGTAGAATCAAGGAGCGGAGAGTCTCAGACAAAGCCTAATCCAGAGCTACCCGGG 189

Qy 94 -----AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107

Db 190 AATTTCTGTGCCAACACAGTGCACGCTGGAGCTCCCGCCAGCTCTCAAGCACTGCTG 249

Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127

Db 250 CTGGGGTGGGTATCCACAGGCTGGTACCTGCCCTCTATGGGCTTGTGGGTGGGG 309

Qy 128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147

Db 310 CTGCTTCCCAATGGGCTGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 369


```
Qy 86 LysLeuLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db 335 -----TGTGCAATGACAGTACACC---CTGGAGTCCCGGACAGTACAGGCA 382
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaLeuTyrLeuValPheVal 125
Db 383 CTGCTTCTGGCTGGTGGCCACAGGCTGGTGGCCCTCTATGGCTGGTCTGGT 442
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 443 GTGGGCTGGCCGCAATGGCTGGCTGGTGGTGGTGGCCAGCAGGACCTGGCTG 502
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaLeuAlaAspPheLeuPheCysValThrLeu 165
Db 503 CCTCCACCATGCTGCTGATGAACCTGGCACTCTGACCTCTCTGCTGGCCCTGGCTG 562
Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db 563 CCCCCGGATCGCTACCACTCGTGGCCAGCGCTGGCCCTTCGGGGAGCGCGCTGC 622
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 623 CGCTGGCCACGGCCGCACTATGGTCAATGATGCTCAGTGTCTGCTGGCCGCC 682
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
Db 683 GTCAGCTGATCGTACTACCTGGCCCTGGTGCACCGCTGGCGGCCCGCCCTGGCGGC 742
Qy 226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db 743 CGCGCGCTGGCCCTGGCACTCTGATGGCTGCTGGCTCATGGCGCGCCCTGGCACTG 802
Qy 246 ProPhePheLeuLysGlnGluTyrLeuValGlnProAspIleThrCysHis 265
Db 803 CCGCTGACATGACGGGAGACCTTCGGCTGGCGCGCTCCGATCGCTGCTGCTGCTG 862
Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db 863 GAGCGCTGCGCCCTGGACGACAGCGCTCCCACTGGCAA---CGGCGCTTCACCTGCCTG 919
Qy 286 AlaPhePheGlyPheLeuLeuProPheValLeuIleLeuTyrCysTyrAlaAlaIle 305
Db 920 GCGCTGTGGCTGTGTTCTGGCCCTCTGCTGGCCATGCTGCTGCTACGGGGCCACCTG 979
Qy 306 ArgThrLeuAsnAlaTyrAspHisArgThrLeuTyrTyrValLysAlaSerLeuLeu 325
Db 980 CACAGCTGGCGGCGGCGCGCTACGGCCACCGCTGAGCTGAGCTGAGCTGAGTGTG 1039
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleLeuLeuIleHisAla 345
Db 1040 CTGGCTCCGCGGTGGCTTCTTCTGTCGCCAGCAACCTGCTGCTGCTGCTGCTGCTG 1099
Qy 346 AsnTyrTyrTyrAsnAnthrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db 1100 GACCGAGCCCGGCGCTGGGCAACCTCTATGTTGCTACGTCAGTCCGACGCTGGCTG 1159
Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1160 AGCACCTCAACAGCTCGGTGGATCCCTTCATCTACTACTACTGCTGCTG 1207

RESULT 22
AAZ50775
ID AAZ50775 standard; DNA; 4895 BP.
XX
AC AAZ50775;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human protease activated receptor-4 DNA.
XX
KW Human; PAR-4; protease activated receptor;
KW antisenase molecule; PAR antibody; cytostatic; therapeutic;
```

```
KW metastatic tumour cell; placental implantation; invasive cell; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 176..1333
FT /tag= a
FT /product= "PAR-4"
XX
XX WO200008150-A1.
XX
XX 17-FEB-2000.
XX
XX 05-FEB-1999; 99WO-IL00079.
XX
XX 07-AUG-1998; 98IL-0125698.
XX
XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
XX
XX Bar-Shavit R;
XX
XX WPI; 2000-205706/18.
XX
XX P-PSDB; AAY45036.
XX
XX Treating metastatic tumor cells useful for treating disorders involving
XX placenta implantation in a female comprises administration of an
XX antisenase molecule complementary to an RNA sequence of a protease
XX activated receptor protein -
XX
XX Example 3; Fig 11a; 46pp; English.
XX
XX The patent discloses a method to treat metastatic tumor cells using
XX an antisenase molecule comprising a polynucleotide complementary to an
XX RNA sequence of a protease activated receptor (PAR) protein, or an
XX antibody capable of binding to a PAR protein. The antisenase molecules
XX and antibodies of PAR protein are also used to treat disorders associated
XX with implantation of placenta. The present sequence is a human PAR-4
XX DNA used for producing antisenase molecules for treating invasive
XX cells.
XX
XX Sequence 4895 BP; 971 A; 1573 C; 1400 G; 951 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 3,92e-38 Length: 4895
XX Score: 523.50 Matches: 121
XX Percent Similarity: 53.57% Conservative: 59
XX Best Local Similarity: 36.01% Mismatches: 145
XX Query Match: 24.27% Indels: 12
XX DB: 21 Gaps: 5
XX
XX US-09-208-629F-6 (1-408) x AAZ50775 (1-4895)
Qy 49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
Db 225 CTGGCGGACCCAGACCCCGGCTACGACGAGCGGAGCGGACCGGA-GGTGGTGAT 283
Qy 68 GluGluPhePro-----PheSerAlaLeuGluGlyTyrThrClyAlaThrIleThrVal 85
Db 284 GACAGCACCGCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 334
Qy 86 LysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly 105
Db 335 -----TGTGCAATGACAGTACACC---CTGGAGTCCCGGACAGTACAGGCA 382
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaLeuTyrLeuValPheVal 125
Db 383 CTGCTTCTGGCTGGTGGCCACAGGCTGGTGGCCCTCTATGGCTGGTCTGGT 442
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 443 GTGGGCTGGCCGCAATGGCTGGCTGGTGGTGGTGGCCAGCAGGACCTGGCTG 502
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
```

```
Db 503 CCCTCCACCATGCTGCTGATGAACCTCGCGACTGCTGACCTCTGCTGGCCCTGGCGCTG 562
Qy 166 PropHelysileAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db 563 CCCCGCGGATCGCTACCACTCGCTGCGTGGCGACGCTGGCCCTTCGGGGAGCGCGCTGC 622
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 623 CGCTGGCCACGGCCGCACTTATGTGTCACATGATGCTCAGTGTCTGCTGGCGGCC 682
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
Db 683 GTCAGCTGATCGTACCTGGCCCTGGTGCACCCGCTGGCGGCCCGCGCTGGTGGC 742
Qy 226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
Db 743 CGCGCGCTGGCCCTGGCACTTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 802
Qy 246 PropHelysileLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db 803 CCCTGACACTGACGGCGAGACCTTCGGCTGGCGGCTCCGATCGCTGCTCTGCCAT 862
Qy 266 AspValHisAsnThrCysGluSerSerPropHelysileLeuTyrTyrPheIleSerLeu 285
Db 863 GACGCGCTGCCCTGGAGCGCACAGGCTCCCACTGGCAA---CGGCGCTTCACCTGCCTG 919
Qy 286 AlaPhePheGlyPheLeuIleProPheValLeuIleTyrCystyrAlaAlaIle 305
Db 920 GCGCTGTGTGGCTGTTCTTCCTGCCCTGCTGCCATGCTGCTGCTGCTGCTGCTGCTG 979
Qy 306 ArgThrLeuAsnAlaTyrAspHisArgTyrLeuTyrTrpTyrValLysAlaSerLeuIle 325
Db 980 CACAGCTGGCGGCGCGCTACGGCGCTACGGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 1039
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHisAla 345
Db 1040 CTGGCTCCCGCGGCTGCTTCTTCGTCGCCAGCAACCTGCTGCTGCTGCTGCTGCTGCTG 1099
Qy 346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db 1100 GACCGAGCCCGCGCGCTGGCGCACTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1159
Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1160 AGCACCTCACAGCTCGGTGGATCCCTTCTACTACTACTACTACTACTACTACTACTACT 1207

RESULT 23
ID AAH43632 standard; cDNA; 4925 BP.
XX AC AAH43632;
XX DT 21-JAN-2002 (first entry)
XX DE Human PAR4 coding sequence.
XX KW Human; protease-activated receptor; PAR4; N-terminal; activated PAR4;
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 191..1348
XX FT /*tag= a
XX FT /product= "PAR4"
XX FT /note= "CDS derived from protein AAB47623"
XX DN WO200158930-A1.
XX PD 16-AUG-2001.
XX PF 06-FEB-2001; 2001WO-US03807.
```

```
XX 09-FEB-2000; 2000US-0500646.
XX (ZYMO ) ZYMOGENETICS INC.
XX PA Baindur N, West RR;
XX PI WPI; 2001-656678/75.
XX DR P-PSDB; AAB47623.
XX PT Peptides comprising an amino acid sequence are capable of stimulating
XX PT protease-activated receptor 4 which are useful in diagnosis and therapy
XX PT e.g. inhibiting tumor cell proliferation and stimulating platelet
XX PT aggregation .
XX PS Disclosure; Page 51-58; 84pp; English.
XX CC This sequence encodes human protease-activated receptor, PAR4. Peptides
XX CC derived from, or based on, the N-terminal of activated PAR4 (see
XX CC AAB47624-77), are capable of stimulating PAR4. These peptides may be
XX CC used to activate PAR4 at lower concentrations than wild type PAR4.
XX CC These peptides may be used to stimulate platelet aggregation, and for
XX CC inhibiting tumor cell proliferation.
XX SQ Sequence 4925 BP; 977 A; 1582 C; 1411 G; 955 T; 0 other;

Alignment Scores:
Pred. No.: 3,958-38 Length: 4925
Score: 523.50 Matches: 121
Percent Similarity: 53.57% Conservative: 59
Best Local Similarity: 36.01% Mismatches: 145
Query Match: 24.27% Indels: 12
DB: 22 Gaps: 5

US-09-208-629F-6 (1-408) x AAH43632 (1-4925)
Qy 49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
Db 240 CTGGCGGACCCAGACCCCGCTACGACGAGGCGGAGCAGCGCA-GGTGTGTAT 298
Qy 68 GluGluPhePro-----PheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrVal 85
Db 299 GACAGCACGCTCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 349
Qy 86 LysIleLysCysProGluSerAlaSerHisLeuHisValLysValAsnAlaThrMetGly 105
Db 350 -----TGTCGCAATGACAGTACACCC---CTGGAGCTCCCGACAGTACACGGGCA 397
Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheVal 125
Db 398 CTGCTTCTGGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 457
Qy 126 ValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 458 GTGGGGCTGCGGCCAATGGGCTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 517
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db 518 CCTCCACCATGCTGCTGATGACCTCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 577
Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db 578 CCCCGCGGATCGCTACCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 637
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 638 CGCTGGCCACGGCCGCACTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
Db 698 GTCAGCTGATCGCTACCTGGCCCTGGTGGCGGCCCGCGCTGGCGCTGCTGCTGCTGCTG 757
Qy 226 HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu 245
```

```

Db 758 CGGCGCTGGCCCTGGACTGTGATGGCTGTGGCTCAATGGCGCGCCCTGGCACTG 817
Qy 246 ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsnIleThrThrCysHis 265
Db 818 CCCTGACACTGCAGCGGACAGACCTTCGGCTGGCGGCTCCGATCGCTGCTGCAT 877
Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db 878 GAGCGCTGCGCCCTGGACGACACAGCGCTCCCACTGGCAA---CCGGCTTCACTGCGCTG 934
Qy 286 AlaPhePheGlyPheLeuIleProPheValLeuIleTyrCysTyrAlaAlaIle 305
Db 935 GCGCTGTGGCTGTTTCCTGCGCCCTGCTGCGCATGCTGTGTGCTACGGGGCCACCTG 994
Qy 306 ArgThrLeuAsnAlaTyrAspHisArgThrLeuTyrTyrValLysAlaSerLeuIle 325
Db 995 CACAGCTGGCGCCAGCGCGCTGACGGCCACCGCTGAGGTGACCGCAGTGGTG 1054
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHisAla 345
Db 1055 CTGCGCTCCCGCGTGGCTTCTTCGTCGCCAGCAACCTGTGCTGTGCTGCTGCTACTCG 1114
Qy 346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeu 365
Db 1115 GACCGAGCGCCAGCGCTGGGCAACCTCTATGTTGCTGCTACGTGCTGCGCTGGCGCTG 1174
Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 1175 AGCACCTCAACAGCTCGTGGATCCCTTCTACTACTACTACTACTACTACTACTACT 1222

RESULT 24
ID AAX90981 standard; cDNA; 1534 BP.
AC AAX90981;
XX
XX
XX 25-JAN-2000 (first entry)
DE Human protease-activated receptor 4 (PAR4) cDNA.
XX
KW Human protease-activated receptor 4 cDNA; PAR4; thrombin receptor;
KW phosphoinositide hydrolysis; calcium efflux; platelet aggregation;
KW platelet activation; thrombin agonist; thrombin antagonist; therapeutic;
KW wound; blood coagulation; heart attack; stroke; inflammatory response;
KW proliferative response; atherosclerosis; restenosis; glomerulosclerosis;
KW pulmonary inflammation; ARDS; adult respiratory distress syndrome; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 3..1160
XX /*tag= a
XX /*product= "PAR4"
XX
XX WO9943809-A2.
XX
XX 02-SEP-1999.
XX
XX 11-FEB-1999; 99WO-US02983.
XX
XX 27-FEB-1998; 98US-0032397.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Coughlin SR, Kahn M;
XX
XX WPT; 1999-619953/53.
XX
XX P-PSDE; AAY15082.
XX
XX DNA molecules encoding protease-activated receptor 4, useful in
XX compound assays for thrombin agonist and antagonist activity -
XX

```

PS Claim 3; Fig 4; 69pp; English.

XX The present sequence is a cDNA encoding human protease-activated
 CC receptor 4 (PAR4) which is a thrombin receptor expressed on cell surface.
 CC PAR4 is activated by thrombin and mediates signalling events e.g.
 CC phosphoinositide hydrolysis, calcium efflux and platelet aggregation.
 CC The receptor is highly expressed in spleen cells and likely to be
 CC involved in thrombin-mediated activation of platelets and other
 CC haematopoietic cells. It is used for screening novel thrombin agonists
 CC or antagonists. The agonists are used as therapeutics to treat wounds,
 CC promote clotting and as reagents to activate platelets in diagnostic
 CC tests. Antagonists are used to control blood coagulation, treat heart
 CC attacks and strokes, and block inflammatory and proliferative responses
 CC that occur in normal wound healing and variety of diseases including
 CC atherosclerosis, restenosis, pulmonary inflammation (ARDS) and
 CC glomerulosclerosis.

XX Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:

Pred. No.:	9.75e-39	Length:	1534
Score:	523.00	Matches:	121
Percent Similarity:	53.57%	Conservative:	59
Best Local Similarity:	36.01%	Mismatches:	145
Query Match:	24.25%	Indels:	12
DB:	20	Gaps:	5

US-09-208-629F-6 (1-408) x AAX90981 (1-1534)

Qy	49	LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe	67
Db	52	CTGCGCGGACCCAGACCCCGCGCTACGACGAGAGCGGAGCACCGA-GGTGGTGAT	110
Qy	68	GluGluPhePro-----PheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrVal	85
Db	111	GACAGCAGCGCTCAATCTGCTGCCCGCGGCTACCCAGGCCAAGTC-----	161
Qy	86	LysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGly	105
Db	162	-----TGTCGAATGACAGTGACACC---CTGGAGCTCCCGGACAGCTCACGGCA	209
Qy	106	TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheVal	125
Db	210	CTGCTCTCTGGGTGGTGCCACAGGCTGGTGGCGCCCTCTATGGGTGGTCTGGTG	269
Qy	126	ValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIle	145
Db	270	GTGGGCTGCGCGCAATGGCTGGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	329
Qy	146	CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu	165
Db	330	CCCTCCACCATGCTGCTGATGAACCTCGGACTGCTGCTGCTGCTGCTGCTGCTGCTG	389
Qy	166	ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys	185
Db	390	CCCCCGGATCGCTTACCACTCGCTGGCGGCTGGCGGCTGGCGGCTGGCGGCTGGCG	449
Qy	186	ArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys	205
Db	450	CGCTGGCAGCGCCGCACTTATGTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	509
Qy	206	IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys	225
Db	510	GTCAGCTGATCGCTTACCTGGCGCTGGTGGCGGCTGGCGGCTGGCGGCTGGCGGCTG	569
Qy	226	HisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeu	245
Db	570	CGCGCGCTGGCGCTTGGACTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	629
Qy	246	ProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsnIleThrThrCysHis	265
Db	630	CCCTGACACTGACGCGGACAGACCTTCCGCTGGCGGCTGCCGATCCGCTGCTGCCAT	689

Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyrPheIleSerLeu 285
 Db 690 GAGCGCTGCCCTGGACGACACAGCTCCCACTGGCAA---CGGGCTTCACTGCTG 746
 Qy 286 AlaphePheGlyPheLeuLeuProPheValLeuLeuIleTyrCysTyrAlaAlaIle 305
 Db 747 GCGCTGTGGCTGTTTCTGCGCCCTGCTGGCATGCTGCTACGGGGCCACCTG 806
 Qy 306 ArgThrLeuAsnAlaTyrAspHisAtgTTrpLeuTyrVallyAlaSerLeuLeu 325
 Db 807 CACAGCTGGGGCCAGCGCGCTACGGCCAGCTGAGGTGACCGAGTGGTG 866
 Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleLeuLeuIleHisAla 345
 Db 867 CTGGCTCCCGCTGGCTTCTTCTGTCGCCAGCACTGCTGCTGCTGCTACTCG 926
 Qy 346 AsnTyrTyrTrpAsnAsnThrAspGlyLeuTyrPheIleTyrLeuLeuAlaLeuCysLeu 365
 Db 927 GACCCGAGCCCCAGCGCTGGGGCACTCTATGCTGCTACGTGCCAGCTGGCGCTG 986
 Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
 Db 987 AGCACCTCAACAGCTCGGTGGATCCCTTCTACTACTACTACTACTGCTG 1034

RESULT 25

AAF21431
 ID AAF21431 standard; DNA; 1534 BP.

AC AAF21431;

DT 14-MAR-2001 (first entry)

XX Human low adenosine antisense oligonucleotide related sequence #2998.

KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.

XX Homo sapiens.

OS WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -

XX Disclosure; Page 1415; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.

CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.

XX
 SQ Sequence 1534 BP; 235 A; 526 C; 479 G; 293 T; 1 other;

Alignment Scores:

Pred. No.: 9,758-39 Length: 1534
 Score: 523.00 Matches: 121
 Percent Similarity: 53.57% Conservative: 59
 Best Local Similarity: 36.01% Mismatches: 145
 Query Match: 24,25% Indels: 12
 DB: 21 Gaps: 5

US-09-208-629F-6 (1-408) x AAF21431 (1-1534)

Qy 49 LeuAlaLysProThrLeuProIleLysThrPheArg---GlyAlaProProAsnSerPhe 67
 Db 52 CTGGCGGACCCAGACCCCGCTACGACGAGCGGGAGCACCAGGA-GGTGGTGAT 110
 Qy 68 GluGluPhePro-----PheSerAlaLeuGluGlyTTrpThrGlyAlaThrIleThrVal 85
 Db 111 GACAGCAGCCCTCAATCTGCTGCCCCCGCGCTACCCAGCCCAAGTC----- 161
 Qy 86 LysIleLysCysProGluSerAlaSerHisLeuHisValHisValHisAsnAlaThrMetGly 105
 Db 162 -----TGTCCTCAATGACAGTGCACACC---CTGGAGCTCCCGACAGCTCACGGGCA 209
 Qy 106 TyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheVal 125
 Db 210 CTGCTTCTGGGCTGGGTGCCACAGGTGGTGGCCGCTCTATGGGCTGTCTCTGGTG 269
 Qy 126 ValGlyValProAlaAsnAlaValThrLeuTTrpMetLeuPhePheArgThrArgSerIle 145
 Db 270 GTGGGGCTGCGGCCAATGGGTGGCGCTGTGGGTGCTGGCCAGCAGCACCTCGGCTG 329
 Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
 Db 330 CCTCCACCATGCTGCTGATGAACCTCGGACTGTGACCTCTGCTGCGCCCTGGCGCTG 389
 Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTTrpValPheGlyGluValLeuCys 185
 Db 390 CCCCCTGGGATCGCTACACCTGCTGGTGGCAGCGCTGGCCCTTCGGGGAGGCGCGCTGC 449
 Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
 Db 450 CGCCTGGCCACGCGCCACTCTATGCTACATGATGCTGCTGCTGCTGCTGCTGCGGCC 509
 Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLys 225
 Db 510 GTACGCTGGATCGCTACCTGGCCCTGGTGGTGCACCCGCTGCGGGCCGCGCGCTGCTG 569

Db 510 GTCAGCGTGCATCGCTACCTGGCCCTGGTGCACCCGCTGGCGGCCCGCCCTGGCGTGGC 569
Qy 226 HistThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeuTyrMetLeu 245
Db 570 CGCGCGCTGGCCCTTGGACTCTGCATGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629
Qy 246 PhePhePheLeuLeuLeuGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHis 265
Db 630 CCCTGACACTGACGGGAGACCTTCGGCTGGCGGCTCCGATCGCTGCTGCTGCTGCTG 689
Qy 266 AspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeu 285
Db 690 GACGCGCTGGCCCTGGACGACAGCCCTCCCACTGGCAA---CGGCGCTTCACCTGCTG 746
Qy 286 AlaPhePheGlyPheLeuLeuProPheValLeuLeuLeuTyrCysTyrAlaAlaIle 305
Db 747 GCCTGTTGGCTGTTTCTGCGCCCTGCTGGCCATGCTGCTGCTGCTGCTGCTGCTGCTG 806
Qy 306 ArgThrLeuAsnAlaTyrAspHisArgTyrPheLeuTyrValHisAlaSerLeuLeu 325
Db 807 CACAGCTGGCGCCGACGCGCGCTACGGCCAGCGCTGAGGCTGACCGCAGTGGTG 866
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleLeuLeuLeuLeuHisAla 345
Db 867 CTGCGCTCCGCGGCGCTTCTGCTGCCAGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 926
Qy 346 AsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuLeuAlaLeuCysLeu 365
Db 927 GACCCGAGCCCGACGCGCTGGGCAACCTCTATGGTGGCTGCTGCTGCTGCTGCTGCTG 986
Qy 366 GlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
Db 987 AGCACCTCAACAGCTCGGCGATCCCTTCTACTACTACTACTACTACTACTACTACTACT 1034

RESULT 27

AAZ32191
ID AAZ32191 standard; cDNA; 3472 BP.
AC AAZ32191;
XX
DT 13-JAN-2000 (first entry)
XX
DE Human thrombin receptor nucleotide sequence.
XX
KW Human; coding sequence polymorphism; vascular pathology gene;
KW polymorphic site; phenotype correlation; forensic; paternity testing;
KW medicine; genetic analysis; vascular disease; ds.
XX
OS Homo sapiens.
XX
XX W09590454-A2.
XX
XX 07-OCT-1999.
XX
XX 26-MAR-1999; 99WO-US06473.
XX
XX 01-APR-1998; 98US-0054272.
XX
XX (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Lander ES, Daley GO, Cargill M, Ireland JS, Rozen SG;
XX
XX WPI; 1999-620066/53.
XX
XX P-PSDB; AAY49570.
XX
XX Determination of polymorphisms in genes, especially those identifying
XX predisposition to vascular disease
XX
XX Claim 1; Fig 35; 134pp; English.
XX
XX AAZ32159 to AAZ32194 represent reference alleles for specifically
XX claimed nucleic acid sequences from the present invention which comprise
XX polymorphic sites as given in a table in the specification, selected

CC from 92 single nucleotide polymorphisms in which the nucleotide at the
CC polymorphic site is different from a nucleotide at the same site in a
CC reference allele. The nucleic acids, and primers and probes, are used to
CC identify polymorphisms, which may predispose an individual to disease,
CC especially a vascular disease. They can also be used in phenotypic
CC correlations, for forensics, paternity testing, medicine or genetic
CC analysis. AAY49550 to AAY49573 represent the proteins which correspond
CC to some of the reference alleles.

XX Sequence 3472 BP; 933 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:

Pred. No.: 2,8e-38 Length: 3472
Score: 523.00 Matches: 127
Percent Similarity: 52.59% Conservative: 76
Best Local Similarity: 32.90% Mismatches: 153
Query Match: 24.25% Indels: 30
DB: 20 Gaps: 11

US-09-208-629F-6 (1-408) x AAZ32191 (1-3472)

Qy 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
Db 228 GGCGCGCGCGCTGCTGCTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 287
Qy 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db 288 GCCCGCACCG 335
Qy 54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db 336 TTAGATCCCGCGCTATTTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTGGGAG 395
Qy 73 SerAlaLeuGluGlyTyrThrGlyAlaThr-----IleThrValIleLys 88
Db 396 GATGAGGAGAAAAAATGAAAGTGGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
Qy 89 CysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
Db 456 AGTCCTCTTCAAAACAACCTTCTGCTCACTCACTCAGAGATGCTCCGCGATATTGACC 515
Qy 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheValValGlyVal 128
Db 516 AGCTCTGGCTGACACTTTTGTGCCATCTGTGACACCGAGTGTGTTAGTCAGCCTC 575
Qy 129 ProAlaAsn-----AlaValThrLeuTyrMetLeuPhePheArgThrArgSerIleCys 146
Db 576 CCACTAAACATCATGCCCATCGTGTGTTTCATCTCCTGAAATGAAGGTCAAGAAG---CCG 632
Qy 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 633 CGCGTGTGTACATGCTGCTGACCTGGCCAGCGAGATGCTGCTGTTGTGCTGCTGCTGCT 692
Qy 167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValLeuCysArg 186
Db 693 TTTAAGATCAGCTATTACTTTTCCGCGCAGTGTGCGAGTTGGGCTGTAATTTGTGTCG 752
Qy 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 753 TTCGTCACTGCAGCATTTTACTGTAAACATGTACGCTCTATCTGTGTCATGACAGTCATA 812
Qy 207 SerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
Db 813 AGCATTTGACCGGTTTCTGGCTGTGGTGTATCCCATGCGAGTCCCTCTCTGGCGTACTCTG 872
Qy 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrPheAlaThrValPheLeuTyr 243
Db 873 GGAAGGGCTTCTTC-----ACTTGTCTGGCATCTGGGCTTTGGGCATCGCAGGG 923
Qy 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
Db 924 GTAGTGCCTCTGCTCTCAAGGAGCAAAACCATCCAGGTGCCGGGTGCCCAACATCACTACC 983

Qy 264 CyHisaspValHisAasnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
 Db 984 TGTATGATGTCTCAATGAACCTGCTCGAAGGC---TACTATGCTACTACTCTCA 1040
 Qy 284 SerLeuAlaPhePheGlyLeuLeuProPheValLeuLeuLeuTyrCysTyrAlaAla 303
 Db 1041 GCCTTCTCTGCTCTCTCTTTTGTGCGCGTGAATCAATTCACGGTCTGTATGTCT 1100
 Qy 304 IleIleArgThrLeuAasn-----AlaTyrAspHisArgTrpLeu 316
 Db 1101 ATCATTCGATGCTCTAGCTCTCCGAGTTGCCAACCGCAGCAAGAGTCCCGGCTTG 1160
 Qy 317 TrpTyrVallyAlaSerLeuLeuLeuVallePheThrLeCysPheAlaProSer 336
 Db 1161 TTC-----CTGTGAGCTGTCTTTTCTGTCATCTTCATCTTCTGCGACCCACA 1211
 Qy 337 AsnIleLeuLeuIleHisAlaAasnTyr---TyrTyrAasnThrAspGlyLeu 355
 Db 1212 AAGTCTCTCTGATGCGCATTAATCTCTTCTCACACTTCCACACAGAGGTGCC 1271
 Qy 356 TyrPheIleTyrLeuLeuAlaLeuCysLeuGlySerLeuAasnSerCysLeuAspProPhe 375
 Db 1272 TACTTTGCTACCTCTCTGCTGTCTGTGTCAGCAGCATTAAGCTGCTGTCATGCCCTTA 1331
 Qy 376 LeuTyrPheLeuMetSer 381
 Db 1332 ATTACTATTACGCTTCC 1349
 RESULT 28
 AAZ50771
 ID - AAZ50771 standard; DNA; 3480 BP.
 AC AAZ50771;
 XX
 XX 31-MAY-2000 (first entry)
 DT Human thrombin receptor DNA.
 DE Human; thrombin receptor; Thr; PAR-1; protease activated receptor;
 KW antisease molecule; PAR antibody; cytostatic; therapeutic;
 KW metastatic tumour cell; placental implantation; invasive cell; ds.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT CDS 225..1502
 FT /*tag= a
 FT /product= "Thrombin receptor"
 FT /transl_except= (pos:1500..1502, aa:Glx)
 FT /note= "no stop codon given"
 FT misc_binding 320..570
 FT /*tag= b
 FT /bound_moiety= "Thr RNA probe"
 XX
 PN W0200008150-A1.
 XX
 XX 17-FEB-2000.
 XX
 XX 05-FEB-1999; 99WO-IL00079.
 XX
 XX 07-AUG-1998; 98IL-0125698.
 XX
 XX (HADA-) HADASIT MEDICAL RES SERVICES & DEV.
 XX
 XX Bar-Shavit R;
 XX
 XX WPI; 2000-205706/18.
 DR P-PSDB; AAY45035.
 XX
 XX Treating metastatic tumor cells useful for treating disorders involving
 PT placenta implantation in a female comprises administration of an
 PT antisease molecule complementary to an RNA sequence of a protease
 PT activated receptor protein -

XX Example 2; Fig 1a; 46pp; English.

PS The patent discloses a method to treat metastatic tumour cells using
 CC an antisease molecule comprising a polynucleotide complementary to an
 CC RNA sequence of a protease activated receptor (PAR) protein, or an
 CC antibody capable of binding to a PAR protein. The antisease molecules
 CC and antibodies of PAR protein are also used to treat disorders associated
 CC with implantation of placenta. The present sequence is a DNA encoding
 CC human thrombin receptor (Thr) (also known as PAR-1), which is a
 CC G-coupled protein belonging to the PAR family. This sequence is used to
 CC produce antisease molecules or antibodies against PAR protein for
 CC therapeutic use.

XX Sequence 3480 BP; 941 A; 817 C; 785 G; 937 T; 0 other;

Alignment Scores:
 Pred. No.: 2,81e-38 Length: 3480
 Score: 523.00 Matches: 127
 Percent Similarity: 52.59% Conservative: 76
 Best Local Similarity: 32.90% Mismatches: 153
 Query Match: 24.25% Indels: 30
 DB: 21 Gaps: 11

US-09-208-629F-6 (1-408) x AAZ50771 (1-3480)

Qy 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
 Db 228 GGGCGGGGGGCTGCTGTGGCGGCTGCTTTCAGTGTGGCGCGCTGTGTCT 287
 Qy 34 LeuProThrPheCysGlnSerGlyMetGluAasnAspThrAsnAsnLeuAlaLysProThr 53
 Db 288 GCCCGCACCGGGCGCGGCGGAGCAATCAAAAGCAACAAT-----GCCACC 335
 Qy 54 LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
 Db 336 TTAGATCCCGGTCATTTCTTCTCAGGAACCCCAATGATAAATATGAACATTTGGGAG 395
 Qy 73 SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrVallyIleLys 88
 Db 396 GATCAGGAGAAATGAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
 Qy 89 CysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThr 108
 Db 456 AGTCTCTTCAAAACAACTTCTGTCATCTCAGAAAGATCGCTCCGATATTGACC 515
 Qy 109 SerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGlyVal 128
 Db 516 AGCTCTGCTGACACTCTTTGTCCCATCTGTACACCGGAGTGTGTGTAGTCAGCCTC 575
 Qy 129 ProAlaAasn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
 Db 576 CCACTAAACATCATGCGCATCTGTTGTTCATCTCTGAAATGAAGGTCAAGAAG---CCG 632
 Qy 147 ThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
 Db 633 CGCGTGGTGATCATGCTGCACCTGCCACCGCAGATGTCTGTTGTGTCTGCTCCCC 692
 Qy 167 PheLysIleAlaTyrHisLeuAasnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
 Db 693 TTTAAGATCAGCTATTACTTTTCCCGCAGTGTGCGAGTTGGGCTGTGAATTTGTGTCGC 752
 Qy 187 AlaThrThrValIlePheTyrGlyAasnMetTyrCysSerIleLeuLeuAlaCysIle 206
 Db 753 TTGCTCAGTCAGCATTTTACTGTAACTGATGCGCTCTAATCTGTCTCATGACAGTCATA 812
 Qy 207 SerIleAasnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGlyLeu 223
 Db 813 AGCATTGACCGGTTTCTGGCTGTGGTGTATCCCATGCAGTCCCTCTCTCGCGTACTCTG 872
 Qy 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
 Db 873 GGAAGGGCTTCCTTC-----ACTTGTCTGGCCATCTCTGGGCTTTGGCCATCGCAGGG 923

Db 760 TGTATGATGTCTCAATGAACCCCTGCTCGAAGGC---TACTATGCTACTACTTCTCA 816
 Qy 284 SerLeuAlaPheGlyPheLeuLeuProPheValLeuLeuIleTyrCysTyrAlaAla 303
 Db 817 GCCTTCTCTGTCTCTCTTTTGTGCGCGTGTATCTATTCACCGGTCTGTATGTCT 876
 Qy 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrpLeu 316
 Db 877 ATCATTCGATGTCTTAGCTCTTCGCGAGTTGCGCAACCGCAGCAAGAAAGTCCCGGCTTTG 936
 Qy 317 TrpTyrValIysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSer 336
 Db 937 TTC-----CTGTACGCTGTCTTTCTGCAATCTTCATCATTTGCTTCGGACCCACA 987
 Qy 337 AsnIleIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
 Db 988 AACGCTCCTCTGATTGGCATTACTCATCTCTTCTCACACTTCCACCACAGAGGTGCC 1047
 Qy 356 TyrPheIleTyrIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
 Db 1048 TACTTTGCCCTACCTCTCTGTGTGTGTCAGCAGCATAAAGCTGCTGCATCGACCCCTTA 1107
 Qy 376 LeuTyrPheLeuMetSer 381
 Db 1108 ATTACTATTACGCTTCC 1125

RESULT 30
 ID AAQ73590
 XX AAQ73590 standard; DNA; 1764 BP.
 AC AAQ73590;
 XX
 DT 25-JUN-1995 (first entry)
 XX
 DE Fragment of the human thrombin receptor gene.
 XX
 KW TR; expression; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 376..1653
 FT CDS /*tag= a
 XX
 PN W09421789-A.
 XX
 PD 29-SEP-1994.
 XX
 PF 28-FEB-1994; 94WO-US02388.
 XX
 PR 25-MAR-1993; 93US-0038662.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Betlach MC, Turner GJ;
 XX
 DR WPI; 1994-317010/39.
 DR P-PSDB; AAR60698.
 XX
 PT Expression of heterologous proteins in halo-bacteria - using
 PT regulatory and stop sequences from halo-bacteria, pref. the
 PT bacterio-rhodopsin gene.
 XX
 PS Disclosure; Fig 14; 118pp; English.
 XX
 CC The sequence is that of the human thrombin receptor gene fragment.
 CC The gene is used to exemplify a new expression vector for producing
 CC heterologous polypeptides in a halobacterial host.
 CC See also AAQ73586-92.
 XX
 SQ Sequence 1764 BP; 382 A; 492 C; 421 G; 469 T; 0 other;
 Alignment Scores:

Pred. No.: 4.65e-38 Length: 1764
 Score: 516.50 Matches: 123
 Percent Similarity: 53.05% Conservative: 77
 Best Local Similarity: 32.63% Mismatches: 144
 Query Match: 23.95% Indels: 33
 DB: 15 Gaps: 11

US-09-208-629F-6 (1-408) x AAQ73590 (1-1764)

Qy 30 LeuLeuLeuLeuLeuProThrPheCysGln-----SerGlyMet 42
 Db 376 ATGTTGGAGTTATTGCCACAGCAGTGGAGGGGTATCGCAGGCCAGATCCAGGCGCTG 435
 Qy 43 GluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAla 62
 Db 436 GACTACAAGGACGATGATGACGTGCGCGCCACCTTAGATCCCGGTCATTCTTCTCAGG 495
 Qy 63 ProProAsn---SerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAla 81
 Db 496 AACCCCAATGATAAATATGAACCATTTTGGAGGATGAGGAGAAATGAAGTGGGTTA 555
 Qy 82 Thr-----IleThrValIysIleLysCysProGluGluSerAlaSerHisLeu 97
 Db 556 ACTGAATACAGATTAGTCTCCATCAATAAAGCAGTCTCTTCAAAACAACTTCCTGCA 615
 Qy 98 HisValIysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuPro 117
 Db 616 TTCATCTCAGAAGATGCTCCGGATATTGACCAGCTCCTGGCTGACACTCTTTGTGCCA 675
 Qy 118 AlaIleTyrLeuLeuValPheValGlyValProAlaAsn-----AlaValThrLeu 135
 Db 676 TCTGTGTACACCGGAGTGTGTGTAGTCAGCTCCCAATCAATCATGCGCATCTGTTGTG 735
 Qy 136 TrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsnLeuAla 155
 Db 736 TTCATCTGAAATGAAGTCAAGAAG---CCGCGCGTGTGTACATGCTGACCTGGCC 792
 Qy 156 IleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGly 175
 Db 793 ACGGCAGATGTCTGTGTGTCTGTCTCCCTTTAAGATCAGTATTACTTTTCGGC 852
 Qy 176 AsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsn 195
 Db 853 AGTGATGGCAGTTTGGGTCTGAATTTGTGTGCTCTCATCTGACGACATTTTACTGTAA 912
 Qy 196 MetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaLeuVal 215
 Db 913 ATGTACGCTCTATCTTGTCTGACAGTCATGAGCATTTGACCGGTTCTGGCTGTGGTG 972
 Qy 216 HisPro-----PheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThr 232
 Db 973 TATCCCATGCGAGTCCCTCTCTGCGTACTCTGGGAAGGGCTTCCTTC-----ACT 1023
 Qy 233 CysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGln 252
 Db 1024 TGTCTGGCATCTGGGCTTTGGCCATCGCAGGGGTAGTGCCTCTCGCTCAAGAGCAA 1083
 Qy 253 GluTyrTyrLeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGlu 272
 Db 1084 ACCATCCAGGTGCCCGGCTCAACATCACTACTCTGTCATGATGTGTCATGAACACCTG 1143
 Qy 273 SerSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuLe 292
 Db 1144 CTGGAAGGC---TACTATGCTACTACTTCTCAGCCTTCTCTGCTCTCTCTTTTGTG 1200
 Qy 293 ProPheValLeuIleIleTyrCysTyrAlaAlaIleArgThrLeuAsn----- 309
 Db 1201 CCGCTGATCATTTCCACGGTCTGTATGTCTCATTCATTCGATGCTCTCTCTCGCA 1260
 Qy 310 -----AlaTyrAspHisArgTrpLeuTyrTyrValIysAlaSerHisLeuLeu 325
 Db 1261 GTTGCCAACGGCAGCAAGATCCCGGCTTTGTTTC-----CTGCAGCTGTGTT 1311

```
Qy 326 LeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAla 345
Db 1312 TTCTGATCTTCATCTTCTCGGACCCACAAACGTCCTCTGATTGGCATCTCA 1371
Qy 346 AsnTyr---TyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCys 364
Db 1372 TTCCTTTCTCACACTTCCACACAGAGGCTGCTACTTTGCTACCTCTCTCTGTCTGT 1431
Qy 365 LeuGlySerLeuAsnSerCysLeuAspPheLeuTyrPheLeuMetSer 381
Db 1432 GTCAGCAGCATAAGCTCGTCATCGACCCCTTAATTTACTATTACGCTTCC 1482
RESULT 31
AAT62461
ID AAT62461 standard; DNA; 2910 BP.
XX
AC AAT62461;
XX
DT 16-AUG-1997 (first entry)
XX
DE THR-GPA1 fusion gene.
XX
KW G-protein coupled receptor; agonist; antagonist; assay;
KW G-alpha protein; Gpalp; GPA1 gene; thrombin receptor; ds.
XX
OS Chimaeric Homo sapiens;
OS Chimaeric Saccharomyces cerevisiae.
XX
FH Key Location/Qualifiers
FT 1..2679
FT CDS /*tag= a
FT /product= Thr-GPA1 fusion
FT mRNA 1..39
FT /*tag= b
FT /note= "STET nucleotides 535-573"
FT mRNA 40..1251
FT /*tag= c
FT /product= thrombin receptor
FT /note= "corresponds to nucleotides 288-1499 of
FT the thrombin receptor gene"
FT mRNA 1252..1260
FT /*tag= d
FT /note= "STET nucleotides 1827..1835"
FT mRNA 1261..2910
FT /*tag= e
FT /product= Gpalp
FT /note= "corresponds to nucleotides 202-1851 of
FT GPA1"
FT primer_bind. complement (42..62)
FT /*tag= f
FT /note= "Oligonucleotide g for thrombin receptor
FT gene amplification"
FT primer_bind 1229..1251
FT /*tag= g
FT /note= "Oligonucleotide h for thrombin receptor
FT gene amplification"
XX
FN WO9711159-A1.
XX
PD 27-MAR-1997.
XX
PF 20-SEP-1996; 96WO-US15203.
XX
PR 20-SEP-1995; 95US-0004023.
XX
PA (HEAR-) HEARTLAND BIOTECHNOLOGIES LLC.
XX
PI Boulton TG, Das PO, Mandell RB, McMullen TW;
XX
DR WPI; 1997-202868/18.
DR P-ESDB; AAW16314.
XX
PT New yeast constructs for use in screening assays - which express a
```

```
PT fusion comprising a mammalian or fungal transmembrane receptor and
PT a yeast G-alpha protein
XX
PS Example 2; Fig 12A-G; 109pp; English.
XX
CC A gene fusion (AAT62461) codes for a polypeptide (AAW16314) comprising
CC the human thrombin receptor (Thr) covalently linked to yeast G-alpha
CC protein Gpalp (encoded by the GPA1 gene). It was obtd. by PCR
CC amplification (see also AAT62466-67) of the Thr gene and insertion
CC of the amplified sequence into pRMH33, which contains the GPA1
CC gene. Thrombin-dependent activation of the yeast mating pathway
CC by the Thr-GPA1 fusion protein was observed in ste2 gpal- yeast
CC cells. The polypeptide is an example of novel fusion proteins
CC between eukaryotic heterotrimeric G-protein coupled receptors and
CC yeast G-alpha proteins that can be coupled to the pheromone-
CC induced signal transduction pathway of yeast and used in screening
CC assays to identify agonists or antagonists of the receptor.
XX
SQ Sequence 2910 BP; 868 A; 606 C; 628 G; 808 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,31e-37 Length: 2910
Score: 512.00 Matches: 118
Percent Similarity: 54.47% Conservative: 71
Best Local Similarity: 34.01% Mismatches: 133
Query Match: 23.74% Indels: 26
DB: 18 Gaps: 9
US-09-208-629f-6 (1-408) x AAT62461 (1-2910)
Qy 52: ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
Db 86 CCTTAGATCCCGGTTCATTTCTTCAGGACCCCAAT-GATAAATATGAACCATTTGG 144
Qy 72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr-----lleThrValIle 87
Db 145 GAGGATGAGAGAGAAATGAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAA 204
Qy 88 LysCysProGluSerAlaSerHisLeuHisValIleAsnAlaThrMetGlyTyrLeu 107
Db 205 AGCAGTCTCTTCAAAAACAACCTTCCTGCATTCATCTCAGAGATGCTCCGCGATATTG 264
Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheValValGly 127
Db 265 ACCAGTCTCTGGCTGACACTTTTGTCCCATCTGTGTACACCGAGTGTGTAGTCAGC 324
Qy 128 ValProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
Db 325 CTCACCAATAACATCATGCGCATCTGTGTCTCATCTGAAAAATGAAGGTCAAGAAG--- 381
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db 382 CCGCGGTGGTGATACATGTCACCTGGCCACCGCAGATGTGCTGTGTGTGTCTGTGCTC 441
Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185
Db 442 CCCTTTAAGATCAGCTATTACTTTTCGGCAGTGATTGGCAGTTGGGTCTGAATTGTGT 501
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 502 CGCTTCGTCACATGACGCAATTTACTGTAAATGATAGCCCTCTATCTGTGTCATGACAGTC 561
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGly 222
Db 562 ATAAGCATTTGACCGGTTTCTGGCTGTGGTGATATCCCATGTCAGTCCCTCTCTGGCGTACT 621
Qy 223 LeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242
Db 622 CTGGGAGGCTTCCTTC-----ACTTGTCTGCCCATCTGGGCTTTGGCCATCGCA 672
Qy 243 TyrMetLeuProPhePheLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262
Db 673 GGGGTAGTCTCTCGTCTCTCAAGAGCAAAACCATCCAGGTGCGCGGCTCAACATCACT 732
```


QY 206 IISerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGly 222
 DB 520 ATAAGCATTCACCGTCTTGGCTGTGTATCCCATGCAGTCCCTCTCTGGCGACT 579
 QY 223 LeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeu 242
 DB 580 CTGGGAAGGCTTCCTTC-----ACTTGCTGGCCATCTGGGCTTTGGCCATCGCA 630
 QY 243 TyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThr 262
 DB 631 GGGGTAGTGCCTCTCGCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACT 690
 QY 263 ThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPhe 282
 DB 691 ACCTGTCATGTGTCTCAATGAACCTCTCGAAGGC---TACTATGCTACTACTTC 747
 QY 283 IISerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAla 302
 DB 748 TCAGCTTCTCTGCTGCTCTTTTGTGCGCGTGATCATTTCCACGGTCTGTATGTG 807
 QY 303 AlaIleIleArgThrLeuAsn-----AlaTyrAspHisArgTrp 315
 DB 808 TCTATCATTCGATGTCTTAGCTTTCGCGAGTTGCCAACCGCAGCAAGAGTCCCGGCT 867
 QY 316 LeuTrpTyrValIlyAlaSerLeuLeuIleValIlePheThrIleCysPheAlaPro 335
 DB 868 TTGTTC-----CTGTCAGCTGCTGTTTCTGCACTTCATCATTTCTTCGACCC 918
 QY 336 SerAsnIleIleLeuIleHisHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354
 DB 919 ACAACGCTCTCTGATGGCATTAATCTCTTCTCACATTCACACACAGAGGCT 978
 QY 355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
 DB 979 GCCTACTTTGCTACCTCTCTGTGTCTGTGTGTCAGCAGCATAAAGCTCGTCATCGACCCC 1038
 QY 375 PheLeuTyrPheLeuMetSer 381
 DB 1039 CTAATTACTATTACGCTTCC 1059
 RESULT 33
 AAA35310
 ID AAA35310 standard; DNA; 3182 BP.
 XX
 AC AAA35310;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:184.
 XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytotstatic; analgesic; impaired airway;
 KW lung disease; ischemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009525-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17712.
 XX
 PR 03-AUG-1998; . 98US-0095212.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;

XX WPI; 2000-205971/18.
 XX New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers
 XX
 XX Disclosure; Page 1330-1331; 1343pp; English.
 XX
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytotstatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers which may metastasize to the lungs, including
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA3312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, and then the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32333 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 SQ Sequence 3182 BP; 895 A; 701 C; 684 G; 902 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,6e-37 Length: 3182
 Score: 512.00 Matches: 118
 Percent Similarity: 54.47% Conservative: 71
 Best Local Similarity: 34.01% Mismatches: 133
 Query Match: 23.74% Indels: 26
 DB: 21 Gaps: 9
 US-09-208-629f-6 (1-408) x AAA35310 (1-3182)
 QY 52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
 DB 44 CCTTAGATCCCCGGTTCATTTCTTCAGGACCCCAAT-GATAAATATGAACCATTTGG 102
 QY 72 PheSerAlaLeuGluGlyTrpThrGlyAlaThr-----lleThrValIle 87
 DB 103 GAGGATGAGGAGAAAAATGAAGTGGTTAACTGAATACAGATTAGTCTCCATCAATAAA 162
 QY 88 LysCysProGluSerAlaSerHisLeuHisValIlyAsnAlaThrMetGlyTyrLeu 107
 DB 163 AGCAGTCTCTTCCAAAAAACAACTTCTCGCATTCATCTCAGAGATGCTCCGATATTTG 222
 QY 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheValValGly 127
 DB 223 ACCAGTCTCTGGCTGACACTCTTTGTCCTCATCTGTGTACACCGAGTGTGTGTAGTCAGC 282
 QY 128 ValProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIle 145
 DB 283 CTCCCACTAAACATCATGGCCATCGTGTGTCTTCTCTGAAATGAAGGTCGAAGAG-- 339
 QY 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
 DB 340 CCGCGGTGTGTGTACATGTGTCACCTGGCCACGCGAGATGTCTGTGTGTGTCTGTC 399
 QY 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCys 185

```

Db 400 CCCCTTTAAGATCAGCTATTACTTTTCGGCGAGTGATTGGCAGTTGGTCTGAATTGTT 459
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeuAlaCys 205
Db 460 CGCTTCGTCATCGCAGCATTTTATGCTAAACATGATACGGCTTATCTATCTGCTCATGACAGTC 519
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGly 222
Db 520 ATAAGCATTTGACCGGTTTCTGGCTGTGGTGATGCCATGCCAGTCAGTCCCTCTCTCGCGGTACT 579
Qy 223 LeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeu 242
Db 580 CTGGGAAGGGCTCTCTTC-----ACTTGTCTGCCCATCTCGGGCTTTGGCCATCGCA 630
Qy 243 TyrMetLeuProPheIleLeuLysGlnGluTyrTyrValGlnProAspIleThr 262
Db 631 GGGGTAGTCTCTCGCTCAAGAGCAACCATCCAGTGCCGGGCTCAACATCACT 690
Qy 263 ThrCysHisAspValHisAsnThrCysGluSerSerPropheGlnLeuTyrTyrPhe 282
Db 691 ACCTGTATGATGTGCTCAATGAACCCCTGCTCGAAGGC---TACTATGCCTACTACTTC 747
Qy 283 IleSerLeuAlaPhePheGlyPheIleProPheValLeuIleIleTyrCysTyrAla 302
Db 748 TCAGCCTTCTCTGCTGCTTCTTTTGTGGCGCTGATCATTTCCACGGTCTGTATGTG 807
Qy 303 AlaIleIleArgThrLeuAsn-----AlaTyrAspHisArgTyr 315
Db 808 TCTATCATTCGATGCTTAGCTCTTCGCGAGTTCGCCAACCCGACAGAGTCCCGGCT 867
Qy 316 LeuTyrTyrValLysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaPro 335
Db 868 TTGTTTC-----CTGTACGCTGCTTTTCTGCACTTTCATCATTTGCTTCGGACCC 918
Qy 336 SerAsnIleLeuIleIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGly 354
Db 919 ACAACGCTCTCTGATTCGCGATTACTCATCTCTTCACACTTCCACCCACAGAGGCT 978
Qy 355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db 979 GCCTACTTTGCTACCTCTCTCTGCTGTGTGTCAGCAGCATAGCTCGTGATCGACCCC 1038
Qy 375 PheLeuTyrPheLeuMetSer 381
Db 1039 CTAATTTACTATTACGCTTCC 1059
RESULT 34
ABK70887
ID ABK70887 standard; cDNA; 1209 BP.
XX AC ABK70887;
XX AC
XX DT 15-JUN-2002 (first entry)
XX DE Human cDNA encoding PAR1 type thrombin receptor delta 1-49TR.
XX KW Human; ss; gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;
XX KW inflammatory disease; cell proliferative disease.
XX OS Homo sapiens.
XX FN JP2002010784-A.
XX PD 15-JAN-2002.
XX PP 29-JUN-2000; 2000JP-0196514.
XX PR 29-JUN-2000; 2000JP-0196514.
XX PA (TEIJ ) TEIJIN LTD.
XX DR WPI; 2002-321520/36.

```

DR P-PSDB; ABG35298.

XX An inhibitor of cell growth mediated by thrombin used to treat
PT inflammatory and cell proliferative diseases -

XX Example 1; Page 22-24; 44pp; Japanese.

XX The invention relates to a polypeptide or a compound which can inhibit
CC cell growth caused by thrombin. The polypeptide/compound combines to a
CC specific region of the structure of the PAR1 type human thrombin receptor
CC participating to cell growth. Preferably, the compound contains the
CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trip-(X5) X4, X5 = optional
CC amino acid or peptide sequence). Also included are a modified PAR1 type
CC thrombin receptor gene or its fragment used for obtaining the above
CC polypeptide, a human PAR1 type thrombin receptor protein and its
CC encoding DNA comprising a fully. The polypeptide or the compound is used
CC to treat inflammatory diseases and cell proliferative diseases. The
CC present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
CC modified version).

XX Sequence 1209 BP; 254 A; 330 C; 273 G; 352 T; 0 other;

Alignment Scores:

Pred. No.:	1.94e-37	Length:	1209
Score:	507.50	Matches:	121
Percent Similarity:	52.32%	Conservative:	71
Best Local Similarity:	32.97%	Mismatches:	138
Query Match:	23.53%	Indels:	37
DB:	24	Gaps:	11

US-09-208-629F-6 (1-408) x ABK70887 (1-1209)

```

Qy 32 LeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLys 51
Db 16 CTGCTGCTGTGGTGGCGCGCTGCTTCAGT-----CTGTGCGGC 51
Qy 52 ProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPhePro 71
Db 52 CCCTGTTGTGTCGCCGAC-----CGGGCCGATAAATATGAACCATTTGG 99
Qy 72 PheSerAlaLeuGluGlyTyrThrGlyAlaThr-----lleThrValLysIle 87
Db 100 GAGGATGAGGAGAGAAATGAAGTGGTTAACTAGTACAGATTAGTCTCCATCAATAAA 159
Qy 88 LysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db 160 AGCAGTCCCTTCCAAAAACAACCTTCCTGCAATTCATCTCAGAGATGCTCCGGATATTG 219
Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuValPheValValGly 127
Db 220 ACCAGCTCTGGCTGACACTCTTTGTCCCATCTGTGTACACCGAGTGTGTTAGTCAGC 279
Qy 128 ValProAlaAsn-----AlaValThrLeuTyrMetLeuPhePheArgThrArgSerIle 145
Db 280 CTCCTCACTAAACATCATGGCCATCGTGTGTTCATCTCTGAAATGAAGGTCAAGAAG--- 336
Qy 146 CysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeu 165
Db 337 CCGGGGTGGTGATCATGTGCACCTGGCCACCGCAGATGTGCTGTTGTTGTCGTGCTC 396
Qy 166 ProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValLeuCys 185
Db 397 CCCTTTAAGATCAGTATTACTTTTCCGGCAGTGATGGCAGTTGGTCTGAATGTTGT 456
Qy 186 ArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCys 205
Db 457 CGCTTCGTCATCGCAGCATTTTACTGTAAACATGTACGCTCTATCTTGTCTATGACAGTC 516
Qy 206 IleSerIleAsnArgTyrLeuAlaIleValHisPro-----PheThrTyrArgGly 222
Db 517 ATAAGCATTTGACCGGTTTCTGGCTGTGGTGTATCCCATGTCAGTCCCTCTCTCGCGTACT 576

```

223	LeuProlyHisThrTyrrAlaLeuValThrCysGlyLeuValThrAlaThrValPheLeu	242
577	CTGGGAAGGCTTCCCTTC-----ACTGTCTGGCCATCTGGGCTTGGCCATCGCA	627
243	TyrMetLeuProPhePheIleLeuLysGlnGluTyrrTyrrLeuValGlnProAspIleThr	262
628	GGGGTAGTGCCTCTGCTCCTCAAGGAGCAACCATCCAGGTGCCGGGCTCAACATCACT	687
263	ThrCysHisAspValHisAsnThrCysGlySerSerSerProPheGlnLeuTyrrTyrPhe	282
688	ACCTGTTCATGATGTGCTCAATGAACCCCTGCTCGAAGGC---TACTATGCTCTACTACTTC	744
283	IleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrrCysTyrAla	302
745	TACGCCTTCTGCTGCTCTCTCTTTTGTGCGCGCTGATCATTTCCACGGTCTGTATTGTG	804
303	AlaIleIleArgThrLeuAsn-----AlaTyrrAspHisArgTrrp	315
805	TCTATCATTCGATGTCTTAGCTCTCCGCGAGTTGCCAACCGCAGCAAGAGTCCCGGGCT	864
316	LeuTrrTyrrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaPro	335
865	TTGTTC-----CTGTCAGCTGCTGTTTTCTGCATCTTCATCTTTCCTTCGGACCC	915
336	SerAsnIleIleLeuIleHisAlaAsnTyr---TyrTyrrAsnThrAspGly	354
916	ACAAAGCTCTCCTCGATTGGCCATTACTCATTCCTTCTCACACTTCCACACAGAGGCT	975
355	LeuTyrrPheIleTyrrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro	374
976	GCCTACTTTTGCCTACTCTCTGCTGTGCTGTGTCAGACGAGTAAAGTCTGCATCGACCCC	1035
375	PheLeuTyrrPheLeuMetSer	381
1036	CTAAATTTACTATTACGCTTCC	1056

RESULT 35

AAQ28568
ID AAQ28568 standard; DNA: 3480 BP.

AA
AC AAQ28568;

DT 15-FEB-1993 (first entry)

XX DE Human thrombin receptor gene.

xx
KW Diagnosis; cardiovascular disease; wound healing; restenosis;
KW thrombosis; unstable angina treatment; myocardial infarction;
KW thrombotic; thromboembolytic stroke; ss.

XX Homo sapiens.

XX	key	Location/Qualifiers
FT	CD8	225..1502
FT		/*tag= a
FT	sig_peptide	225..303
FT		/*tag= b
FT	mat_peptide	304..1502
FT		/*tag= c

XX
PN
WO9214750-A.XX
PD 03-SEP-1992

XX
PF
19-FEB-1992: 92WO-IIS01312

XX
PR
19-FEB-1991: 91US-0657769

PR 07-NOV-1991; 91US-0789184.
YY

PA (CORT-) COR THERAPEUTICS INC.
PA (REGC) UNIV CALIFORNIA.

XX
XX
XXXXXXXXXXXXX

PI Coughlin SR, Scarborough RM;

XX	WPI; 1992-316119/38.
DR	P-PSDB; AAR7240.
DR	
XX	
XX	DNA encoding cell surface receptor for thrombin - useful for
PT	determining thrombin in diagnosing e.g. cardiovascular diseases,
PT	also to treat wound healing, restenosis etc.
PT	
XX	
XX	Disclosure; Fig 1; 81pp; English.
PS	
PS	
XX	The sequence is that of the gene coding for human thrombin receptor.
CC	It can be used in the prepn. of diagnostics to determine thrombin
CC	levels in samples, and screening tools for candidate substances which
CC	affect thrombin activity in vivo. Thrombosis may be diagnosed in a
CC	mammal by measuring the presence, absence or amt. of the cleaved
CC	activation peptide of the TR.
CC	
XX	
SQ	Sequence 3480 BP; 947 A; 816 C; 786 G; 931 T; 0 other;

Alignment Scores:

Argument Scores:		
Pred. No.:	1.6e-36	Length:
Score:	504.00	Matches:
Percent Similarity:	52.07%	Conservative:
Best Local Similarity:	32.38%	Mismatches:
Best Local Similarity:	32.38%	Indels:
Query Match:	23.37%	Gaps:
DB:	13	
		3480

US-09-208-629F-6 (1-408) x AAQ28568 (1-3480)

Qy		14	GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
Db		228	GGCGCCGGCGCTGCTGCTGGTGCCTGTTCAGTCTGTGCGCGCGTGTTGCT 287
Qy		34	LeuProThrPheCysGlnSerGlyWetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db		288	GCCCGCACCCGGCGCCGACGAGCAACAAT-----GCCACC 335
Qy		54	LeuProIleLysThrPheArgGlyAlaProProAsn---SerPheGluGluPheProPhe 72
Db		336	TTTAGATCCCCGGTCATTCTTCTCAGGAACCCCAATGATAAATATGAACCATTTTGGGAG 395
Qy		73	SerAlaLeuGluGlyTrpThrGlyAlaThr-----IleThrVallLysIleLys 88
Db		396	GATGAGGAGAAAATGAAGTGGGTAACTGAATACAGATTAGTCTCCATCAATAAAGC 455
Qy		89	CysProGluGluSerAlaSerHisLeuHisVallysAsnAlaThrMetGlyTyrrLeuThr 108
Db		456	AGTCTCTTTAAAAAACACTTCTCTGCATTCTATCTCAGAAGATGCCCTCCGGATATTTGACC 515
Qy		109	SerSerLeuSerThrLysLeuIleProAlaIleTyrrLeuLeuValPheValValGlyVal 128
Db		516	AGCTCTGGTGCACACTCTTTGTCCCCTGTGTGTACACCGAGTGTTTGTATGCACGCTC 575
Qy		129	ProAlaAsn-----AlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCys 146
Db		576	CCAATAAACATCATGGCCATCGTTGTGTTCATCTGAAATGAAGGTCAAGAAG--CCG 632
Qy		147	ThrThrValPheTyrrThrAsnLeuAlaIleAlaaspPheLeuPheCysValThrLeuPro 166
Db		633	CGGTGGTGTACATCTCCACCTGCCACGCGAGATGTGCTGTTGTGCTGTGCTCCCC 692
Qy		167	PheLysIleAlaTyrrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db		693	TTTAAGATCAGCATATTACTTTTTCCGGCAGTGTGGCAGTTTCGGGTCTGAAATTTGTGTCGC 752
Qy		187	AlaThrThrValIlePheTyrrGlyAsnMetTyrrCysSerIleLeuLeuAlaCysIle 206
Db		753	TTCGTCACTGCAGCATTTTACTGTAAATGTACGCCCTCTATCTTGTCTCATGACAGTCATA 812
Qy		207	SerIleAsnArgTyrrLeuAlaIleValHiPro-----PheThyrrArgGlyLeu 223
Db		813	AGCATTACCGGTTTCTGGCTGTGGTGTATCCCATGAGTCCCTCTCTCTGGCTACTCTGT 872

QY 224 ProlyshisThrTyrAlaLeuValThrCysGlyLeuValTrrAlaThrValPheLeuTyr 243
 Db 873 GGAAGGGCTTCCTTC-----ACTTGTGCGCCATCTGGGTTGGCCATCGCAGG 923
 QY 244 MetLeuProPhePheLeuLeuGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
 Db 924 GTAGTGGCTCTCGCTCAAGAGCAACCATCCAGTGGCCGGGCTCAACATCACHACC 983
 QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
 Db 984 TGTATGATGTCTCAATGAACCTGCTCGAAGGC---TACTATGCTACTACTTCTCA 1040
 QY 284 SerLeuAlaPhePheGlyPheLeuLeuProPheValLeuIleIleTyrCysTyrAlaAla 303
 Db 1041 GCCTTCCTGCTGCTTCAAAAAAGTCCGCTGATCATTTCCACGGTCTGTATGTGTCT 1100
 QY 304 IleIleArgThrLeuAsn-----AlaTyrAspHisArgTrrPhe 316
 Db 1101 ATCATTCGATGCTTAGCTCTTCGCGAGTTCGCAACCGCAGCAAGAGTCCCGGGCTTTG 1160
 QY 317 TrpTyrValLysAlaSerLeuLeuValIlePheThrIleCysPheAlaProSer 336
 Db 1161 TTC-----CTGTGAGTGTCTTTCTGTCATCTTCATCATTTGCTTCGGAGCCACA 1211
 QY 337 AsnIleLeuLeuIleHisAlaAsnTyr---TyrTyrAsnAsnThrAspGlyLeu 355
 Db 1212 AACGTCCTCTGATTGGCATTACTCATTTCTTCTCACACTTCCACACACAGAGCTGCC 1271
 QY 356 TyrPheIleTyrLeuLeuAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPhe 375
 Db 1272 TACTTTGCTACCTCTCTGTCGTCTGTGTCAGCAGCATAGCTCGTGATCGACCCCTTA 1331
 QY 376 LeuTyrPheLeuMetSer 381
 Db 1332 ATTTACTATTACGCTTCC 1349

RESULT 36

ABK70888
 ID ABK70888 standard; cDNA; 1116 BP.

XX AC ABK70888;

XX DT 15-JUN-2002 (first entry)

XX DE Human cDNA encoding PAR1 type thrombin receptor delta 1-80.

XX KW Human; ss: gene; PAR1; thrombin receptor; antiinflammatory; cytostatic;
 XX inflammatory disease; cell proliferative disease.

XX OS Homo sapiens.

XX PN JP2002010784-A.

XX PD 15-JAN-2002.

XX PF 29-JUN-2000; 2000JP-0196514.

XX PR 29-JUN-2000; 2000JP-0196514.

XX PA (TEIJ) TEIJIN LTD.

XX DR WPI; 2002-321520/36.

XX DR P-PSDB; ABG35299.

PT An inhibitor of cell growth mediated by thrombin used to treat
 PT inflammatory and cell proliferative diseases -

XX PS Example 1; Page 24-25; 44pp; Japanese.

XX CC The invention relates to a polypeptide or a compound which can inhibit
 CC cell growth caused by thrombin. The polypeptide/compound combines to a
 CC specific region of the structure of PAR1 type human thrombin receptor
 CC participating to cell growth. Preferably, the compound contains the

CC 52nd to the 56th amino acid sequences at the amino end side of PAR1 type
 CC human thrombin receptor ((X4)-Tyr-Glu-Pro-Phe-Trp-(X5) X4, X5 = optional
 CC amino acid or peptide sequence). Also included are a modified PAR1 type
 CC thrombin receptor gene or its fragment used for obtaining the above
 CC polypeptide, a human PAR1 type thrombin receptor protein and its
 CC encoding DNA comprising a fully. The polypeptide or the compound is used
 CC to treat inflammatory diseases and cell proliferative diseases. The
 CC present sequence is a cDNA encoding a PAR1 type thrombin receptor (or a
 CC modified version).

XX SQ Sequence 1116 BP; 218 A; 318 C; 252 G; 328 T; 0 other;

Alignment Scores:

Pred. No.: 4,09e-37 Length: 1116
 Score: 503.50 Matches: 112
 Percent Similarity: 58.01% Conservative: 69
 Best Local Similarity: 35.90% Mismatches: 105
 Query Match: 23.34% Indels: 26
 DB: 24 Gaps: 10

US-09-208-629F-6 (1-408) x ABK70888 (1-1116)

QY 87 IleIysCysProGlu-----GluSerAlaSerHisLeuHisValLys-AsnAlaTh 103
 Db 57 GTTGTCTGCGCCGACCGGCGCCCAAAACAACTTCTGTCATTCTCAGAAAGATGCGTC 116
 QY 103 rMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeu 123
 Db 117 C---GGATATTTGACAGCTCTCTGGCTGACACTCTTTGTCCATCTGTACACCGGAGT 173
 QY 123 lPheValValGlyValProAlaAsn-----AlaValThrLeuTrpMetLeuPheAr 141
 Db 174 GTTTGTAGTCAGCTCCCACTAAACATCATGCGCATCGTTGTGTTCATCTCTGAAATGAA 233
 QY 141 gThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 161
 Db 234 GGTCAAGAAG---CCGGCGGTGTGTATCATGTCGACCTCGGCCACCGCAGATGTGCTGT 290
 QY 161 eCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 181
 Db 291 TGTGTCTGTGCTCCCTTTAAGATCAGCTATTACTTTCCGGCAGTGTATGGCAGTTGG 350
 QY 181 YGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLe 201
 Db 351 GTCTGAATTGTGTGCTTCTGCTACTGACGACATTTTACTGTAAACATGATGACGCTCTATCT 410
 QY 201 uLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisPro-----Ph 218
 Db 411 GCTCATGACAGTCATAAGCATTCACCGGTTTCTGGCTGTGTGTATCCCATGAGTCCCT 470
 QY 218 eThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrrAl 238
 Db 471 CTCTGCGCTACTCTGGGAAGGGCTTCCTTC-----ACTTGTCTGCCATCTCGGC 521
 QY 238 aThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValG 258
 Db 522 TTTGGCCATCGCAGGGGTAGTGCCTCTGCTCTCAAGGACAAACCATCCAGGTGCCCGG 581
 QY 258 nProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPhe 278
 Db 582 GCTCAACATCCTACCTGCTCATGATGCTCAATGAACCCCTGCTCGAAGGC---TACTA 638
 QY 278 nLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIle 298
 Db 639 TGCTACTACTTCTCAGCCCTTCTCTGCTGCTCTTTTGTGCGCGCTGATCATTTCCAC 698
 QY 298 eTyrCysTyrAlaAlaIleIleArgThrLeuAsn-----AlaTyr 311
 Db 699 GGTCTGTATGTGTCTATCATCTTACGTCTTACGTCTCTCCGAGTTCGCCAACCGCAGCAA 758
 QY 311 rAspHisArgTrrPheLeuTrrTyrValLysAlaSerLeuLeuIleValIlePheThrI 331
 Db 759 GAAGTCCCGGGCTTTGTTTC-----CTGTGAGTGTGCTGTTTCTTCTGTCATCTTTCATCAT 809

Db 792 CACGATCTCTCAGTGAACCTGATGCAAGGC-----TTTACTCGTACTACTTC 842
 Qy 283 IleserLeuAlaPhePheGlyPheLeuIleProPheValLeuIleLeuTyrCysTyrAla 302
 Db 843 TCGGCTTCCTCGGCACTCTTCTTCTGTGCGGTGATCGTTCCACGGTCTGCTACAG 902
 Qy 303 AlaIleleArgThrLeuAsn-----AlaTyrAspHisArgTyr 315
 Db 903 TCCATCATCGGTGCTGAGCTCTCGCGGTGTCACACCGGAGCAAGAAAGTCGCGGCT 962
 Qy 316 LeuTyrPyrVallysAlaSerLeuLeuVallePheThrIleCysPheAlaPro 335
 Db 963 TTGTC-----CTGTCTGCGCGGTGTCGATCTTCATGCTGCTGCTGGGCC 1013
 Qy 336 SerAsnIleLeuIleHisAlaAsnTyrTyr-----AsnAsn 351
 Db 1014 ACCAAGCTCTCTGATTGTGCAC-----TACCTTTCTCTCCGACAGTCCTGGT 1064
 Qy 352 ThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
 Db 1065 ACGGAGCAGCCTACTTGTCTACCTCTCTGCTGCTGTGACGAGTGTGCTGCTGC 1124
 Qy 372 LeuAspProPheLeu---TyrPheLeuMetSerIlyThrArgAsnHisSerThrAlaTyr 390
 Db 1125 ATGATCCGTGATTACTACTAGCTCTCTCGAGTGCCAGAGGACCTCTACAGATC 1184
 Qy 391 Leu-----ThrLys*****AsnAspLeuArgGluGlnGlyGln-----Pro 404
 Db 1185 TTGTCGTCAAAGAAAGCTCTGATCCCAACAGTTGCAACAGCCGCGAGCTGATGCCG 1244
 Qy 405 SerGln 406
 Db 1245 AGTAAA 1250

RESULT 38

AAV40372
 ID AAV40372 standard; cDNA; 2588 BP.

AC AAV40372;
 XX

DT 16-OCT-1998 (first entry)
 XX

XX Mouse G-protein coupled receptor 69A08 #1 encoding cDNA.

XX Mouse; CKDLR201.1; chemokine; G-protein coupled receptor; GPCR;
 KW 7 transmembrane receptor; inflammation; asthma; antiviral;
 KW abnormal cell proliferation; regeneration; degeneration; atrophy; ss.
 XX

OS Mus sp.

XX Key Location/Qualifiers

FT CDS 1..1083

FT /*tag= a

FT /product= "G-protein coupled receptor 69A08"

FT misc_feature 158

FT /*tag= b

FT /note= "residues 158, 159 and 276 are probably absent,
 changing the reading frame between those
 positions; sequence are provided in AAV40373 and
 AAW69598"

XX WO9831810-A2.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-US00218.

XX 21-JAN-1997; 97US-0786624.

XX (SCHE) SCHERING CORP.

XX Capone M, Gorman DM, Hedrick JA, Huffine CF, Rossi DL;
 PI Vicari A, Zlotnik A;

XX WPI; 1998-414108/35.
 DR P-PSDB; AAW69597.
 XX
 PT Pure or recombinant chemokine CKDLR201.1 - useful, e.g. for treating
 inflammation and as antiviral agents
 XX
 PS Claim 7; Page 61-64; 77pp; English.
 XX
 CC The present sequence encodes mouse G-protein coupled receptor 69A08 #1
 protein which comprises a plurality of epitopes. Host cells containing
 vectors comprising a nucleotide sequence encoding the protein are used to
 produce recombinant protein. Treatment of a cell, particularly neuron,
 macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR)
 or a receptor responsive to CKDLR201.1 chemokine protein, with an
 (ant)agonist is used to control physiological development, e.g.
 alteration of calcium ion influx, a chemottractant response, morphology,
 phosphoinositide lipid turnover or an antiviral response. Nucleotide
 sequences the CKDLR201.1 protein, or GPCRs, are useful as primers or
 probes, e.g. for detecting and isolating related sequences and for
 expressing antigenic peptides. Antibodies (Ab) directed against the
 CKDLR201.1 protein and GPCRs are used to detect or purify the proteins;
 diagnostically (e.g. for developmental abnormalities); in screening for
 potential drugs; to inhibit chemokine/receptor activation; (when coupled
 to a toxin or radioisotope) for killing specific cells; and to raise
 anti-idiotypic antibodies. CKDLR201.1 protein and GPCRs and compounds
 which bind them can be used to treat inflammation, e.g. asthma; as
 antiviral agents, and to treat abnormal cell proliferation, regeneration,
 degeneration and atrophy. Therapeutic agents are administered orally, by
 CC injection and rectally.
 XX
 SQ Sequence 2588 BP; 462 A; 782 C; 743 G; 601 T; 0 other;

Alignment Scores:

Pred. No.: 3.28e-35 Length: 2588
 Score: 488.00 Matches: 117
 Percent Similarity: 51.45% Conservative: 60
 Best Local Similarity: 34.01% Mismatches: 134
 Query Match: 22.62% Indels: 34
 DB: 19 Gaps: 4

US-09-208-629F-6 (1-408) x AAV40372 (1-2588)

Qy 56 IleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeu 75
 Db 1135 GTAAAGAGTACCAAGGAGCCAT-----
 Qy 76 GluGlyTyrThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer----- 93
 Db 1159 GAAGGCCCTCTGGGTCCC-----ACAGTACAACCTCAGGAGCGCAAGTCTCAGACAAG 1212
 Qy 94 -----AlaSerHisLeuHisVal 99
 Db 1213 CCTAATCCACGAGGTACCCGGGCAAAATCTGTGCCAACACGACAGTGCAGCTGGAGCTC 1272
 Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
 Db 1273 CCGGCCAGCTCTCAAGCACTGTGTGGGGTGGTCCCAAGAGCTGTACTTGGCTC 1332
 Qy 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhe 139
 Db 1333 TATGGCTTGTGTGGCTGTGGGCTGCTGCCAATGGCTGGCGCTGTGGGTGGCTGGCC 1392
 Qy 140 PheArgThrArgSerIleCysThrThrValPheTyrThr-AsnLeuAlaIleAlaAsp 159
 Db 1393 ACAAGGGTGCCACCGCTGCCATCCACCATCTCTCATGGAACCTGGCAGTGGCTGATCT 1452
 Qy 159 eLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpVa 179
 Db 1453 GCTGTGGCCCTGGTGTGCTGCCACCGACTGGCTTACCACCTTGGCTGGCGCAGCGCTGGCC 1512
 Qy 179 lPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSe 199

Db 1513 ATTGTGTGGAGTGCCTCCGGGTGGCCACAGCTGCCCTCTATGGCCACATGATGGTTC 1572
 Qy 199 rleLeuLeuLeuAlaCysIleSerIleAenArgTyrLeuAlaIleValHisProPheH 219
 Db 1573 AGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1632
 Qy 219 rTyrAtgGlyLeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaH 239
 Db 1633 GGCCCGTGCACCTGGTGTCAACGCCTCACTACTGGACTCTGTTGGTGGCTGGCTCTC 1692
 Qy 239 rValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlyTyrTyrLeuValGlnPr 259
 Db 1693 TGCAGCCACCTGGCTTGCCTCTCACTCTGCATCGCAGAACTCCGATTACTGGCTCC 1752
 Qy 259 oApeIleThrCysHisAspValHisAenThrCysGluSerSerSerProPheGlnLe 279
 Db 1753 GAT-CGATGCTGTGTATGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1811
 Qy 279 uTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIle 299
 Db 1812 G---GCCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1868
 Qy 299 yrCysTyrAlaAlaIleArgThrLeuAenAlaTyrAspHisArgTyrLeuTyrTrpV 319
 Db 1869 TGTGCTATGAACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1928
 Qy 319 allysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleI 339
 Db 1929 TCAGACTGACAGCCCTGGTACTGTCGCGAGTGGCTTCTTTCACACCTTAGCAATGTGC 1988
 Qy 339 leLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAenThrAspGlyLeuTyrPheIle 359
 Db 1989 TGCTGGTGTGCTACTATCAACCCGAGCCCTGAGCCCTGGGCAATCTCTATGGAGCT 2048
 Qy 359 yrLeuIleAlaLeuCysLeuGlySerLeuAenSerCysLeuAspProPheLeuTyrPheL 379
 Db 2049 ATGTGCCAGCTGGCACTCAGACCCCTCAACAGCTGCGTAGACCTTTTCATCTACTACT 2108
 Qy 379 euMetSer 381
 Db 2109 ATGTGTCC 2116
 RESULT 39
 AAV40373
 ID AAV40373 standard; cDNA; 1080 BP.
 AC AAV40373;
 XX
 DT 16-OCT-1998 (first entry)
 XX
 DE Mouse G-protein coupled receptor 69A08 #2 encoding cDNA.
 XX
 KW Mouse; CXDLR201.1; chemokine; G-protein coupled receptor; GPCR;
 KW 7 transmembrane receptor; inflammation; asthma; antiviral;
 KW abnormal cell proliferation; regeneration; degeneration; atrophy; ss.
 XX
 OS Mus sp.
 FH Key Location/Qualifiers
 CDS 1..1080
 FT /*tag= a
 FT /product= "G-protein coupled receptor 69A08"
 XX
 FN MO9831810-A2.
 XX
 PD 23-JUL-1998.
 XX
 PF 20-JAN-1998; 98WO-US00218.
 XX
 PR 21-JAN-1997; 97US-0786624.
 XX
 PA (SCHE) SCHERING CORP.
 XX

PI Capone M, Gorman DM, Hedrick JA, Huffine CF, Rossi DL;
 PI Vicari A, Zlotnik A;
 XX
 DR WPI: 1998-414108/35.
 DR P-PSDB; AAW69598.
 XX
 PT Pure or recombinant chemokine CXDLR201.1 - useful, e.g. for treating
 PT inflammation and as antiviral agents
 XX
 PS Claim 7; Page 66-68; 77pp; English.
 XX
 CC The present sequence encodes mouse G-protein coupled receptor 69A08 #2
 CC protein which comprises a plurality of epitopes. Host cells containing
 CC vectors comprising a nucleotide sequence encoding the protein are used
 CC to produce recombinant protein. Treatment of a cell, particularly neuron,
 CC macrophage or lymphocyte, carrying a G-protein coupled receptor (GPCR)
 CC or a receptor responsive to CXDLR201.1 chemokine protein, with an
 CC (ant)agonist is used to control physiological development, e.g.
 CC alteration of calcium ion influx, a chemottractant response, morpho-
 CC logic, e.g. for detecting and isolating related sequences and for
 CC expressing antigenic peptides. Antibodies (Ab) directed against the
 CC CXDLR201.1 protein and GPCRs are used to detect or purify the proteins;
 CC diagnostically (e.g. for developmental abnormalities); in screening for
 CC potential drugs; to inhibit chemokine/receptor activation; (when coupled
 CC to a toxin or radioisotope) for killing specific cells; and to raise
 CC anti-idiotypic antibodies. CXDLR201.1 protein and GPCRs and compounds
 CC which bind them can be used to treat inflammation, e.g. asthma; as
 CC antiviral agents, and to treat abnormal cell proliferation, regeneration,
 CC degeneration and atrophy. Therapeutic agents are administered orally, by
 CC injection and rectally.
 XX
 SQ Sequence 1080 BP; 177 A; 347 C; 307 G; 249 T; 0 other;

Alignment Scores:
 Pred. No.: 1.46e-35 Length: 1080
 Score: 486.50 Matches: 111
 Percent Similarity: 52.96% Conservative: 59
 Best Local Similarity: 34.58% Mismatches: 135
 Query Match: 22.55% Indels: 17
 DB: 19 Gaps: 2

US-09-208-629F-6 (1-408) x AAV40373 (1-1080)

Qy 76 GluGlyTyrThrGlyAlaThrIleThrValLys-IleLysCysProGluSerAla 95
 Db 3 GAAGGCCCTCTGGGTCCACAGTACAACTCAAGAGCCGGAAGTCTCTCAGACAAAGCTAAT 62
 Qy 95 rHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLys-- 114
 Db 63 CCACGAGGTACCCGGGCAAAATTCGTGCCAAGCAGTGCACAGCTGGAGCTCCGGCC 122
 Qy 115 -----LeuIleProAlaIleTyrLeuLe 122
 Db 123 AGCTCTCAAGCACTGCTGGTGGGTGGTCCACAGTGGTACCTGCTCTATGGCT 182
 Qy 122 uValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgH 142
 Db 183 TGTGTGGTGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 242
 Qy 142 rArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPheCy 162
 Db 243 GCCACGCTGCCATCCATCTGCTCATGAACCTGGAGGCTGGTGGTGGTGGTGGTGGTGG 302
 Qy 162 sValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyG 182
 Db 303 CCTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 362
 Qy 182 uValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLe 202
 Db 363 GGCTGCTCCGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 422

QY 202 uLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGI 222
 Db 423 GCTGGCTGCACTAGCTGGACATACCTGGCCCTGGTGCATCTTTGGGGCCCGTGC 482
 QY 222 YLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLe 242
 Db 483 ACTGCTGGTCAAGCCCTCACTACTGGAATCTGTTGGCTGGCTGCTCTCGAGCCAC 542
 QY 242 uTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleTh 262
 Db 543 CCTGGCTTGCCTCTCACTCTGCATCGCAGAACTTCGGATTACTGGCTCGAT-CGCAT 601
 QY 262 rThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPh 282
 Db 602 GCTGTGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 658
 QY 282 eIleSerLeuAlaPhePheGlyPheLeu-IleProPheValLeuIleIleTyrCysTyrA 302
 Db 659 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 718
 QY 302 laAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrTyrValLysAla 322
 Db 719 GAACCACTTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATGTCGTCATG 778
 QY 322 erLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIle 342
 Db 779 CAGCCCTGGTACTGCTTCGCGCAGTGGCTTCTTTCACACCTAGCAATGTCGTCGTCG 838
 QY 342 leHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleA 362
 Db 839 TGCACTATTCAAAACCCGAGCCCTGAGCCCTGGGGCAATCTCTATGGAGCCTATGTGCCA 898
 QY 362 laLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
 Db 899 GCCTGGCACTCAGCACCTCAACAGCTGCGTAGACCTTTCACTACTACTATGTGTC 957

RESULT 40
 AAZ99588
 ID AAZ99588 standard; DNA; 1955 BP.
 AC AAZ99588;
 XX
 DT 03-JUL-2000 (first entry)
 DE cDNA encoding a human 14400 receptor polypeptide.
 KW Human; 14400 receptor; G-protein coupled receptor; signalling pathway;
 KW CD8 T cell; CD4 T cell; CD34+ bone marrow cell; thrombocytopenia;
 KW inflammation; spleen disorder; splenomegaly; lung disorder;
 KW adult respiratory distress syndrome; colon disorder;
 KW bacterial enterocolitis; liver disorder; hepatic injury;
 KW platelet number; precursor T-cell neoplasm; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 288..1367
 FT /tag= a
 FT /product= "14400 receptor"
 FT /transl_except= (pos: 1119, 1121, aa: Thr)
 XX
 PN WO200011170-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 20-AUG-1999; 99WO-US19112.
 XX
 PR 20-AUG-1998; 98US-0137063.
 PR 20-AUG-1999; 99US-0378100.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Gluckemann MA, Weich NS;

XX WPI; 2000-256378/22.
 DR P-PSDB; AAY69485.
 XX
 PT G-protein coupled receptor, 14400 receptor, for treating CD4, CD8 or
 CD34+ related disorders
 XX
 PS Claim 3; Fig 1A-B; 92pp; English.
 XX
 CC The present sequence encodes a human 14400 receptor polypeptide. The
 CC receptor is a G-protein coupled receptor that participates in signalling
 CC pathways. The polypeptide can be used to identify an agent to modulate
 CC the polypeptides activity in patients having a disorder involving CD8 or
 CC CD4 T cells or CD34+ bone marrow cells, especially where the disorder
 CC involves thrombocytopenia or inflammation. The modulators can
 CC be used to treat disorders involving the spleen (e.g. splenomegaly),
 CC lung (e.g. adult respiratory distress syndrome), colon (e.g. bacterial
 CC enterocolitis), liver (e.g. hepatic injury), the uterus and
 CC endometrium, brain, T-cells, skin, heart, blood vessels, red cells,
 CC thymus, B-cells, kidney, breast, testis and epididymis, prostate,
 CC thyroid, skeletal muscle, pancreas, small intestine, reduced platelet
 CC number, precursor T-cell neoplasms, CD3, CD4, and CD8 T lymphocytes.
 XX
 SQ Sequence 1955 BP; 322 A; 642 C; 567 G; 421 T; 3 other;
 Alignment Scores: Pred. No.: 2,46e-33 Length: 1955
 Score: 466.00 Matches: 104
 Percent Similarity: 51.95% Conservative: 56
 Best Local Similarity: 33.77% Mismatches: 130
 Query Match: 21.60% Indels: 18
 DB: 21 Gaps: 8
 US-09-208-629F-6 (1-408) x AAZ99588 (1-1955)
 QY 90 ProGluGluSerAlaSerHisLeuHisVal-----lysAsnAlaThr 103
 Db 267 CCGGGCGAGGACCCCTCCAGGATGCGAGTCCCGAACAGCAGCCGCGGCAACAGCGACG 326
 QY 104 MetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuVal 123
 Db 327 CTCAGATGTCGGGAACCGGGGATCGCGGTGGCCCTGCCCTGGTGTACTCGCTGGTG 386
 QY 124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArg----- 141
 Db 387 GCGGCGGTGAGCATCCCGGGCAACCTCTCTCTCTGGGTGTGTGCGCGCGCATGGGG 446
 QY 142 ThrArgSerIleCysThrThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeuPhe 161
 Db 447 CCCAGATCC--CCGTCGGTTCATCTTCATGATCAACCTGAGCGTCAGCGACCTGATGCTG 503
 QY 162 CysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGly 181
 Db 504 GCCAGCGTGTGCTTTCCAAATCTACTACCATTCGACCGCCACCGCGGTGTTTCGGG 563
 QY 182 GluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIleLeu 201
 Db 564 GTGCTGCTTTGCAACGTCGTGGTACCGTGGCTTTTACGCAACATGATTCCAGCATCTC 623
 QY 202 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArg 221
 Db 624 ACCATGACCTGTATCAGCGTGGAGCGCTTCTGGGGGTCTCTGTACCGCTGCTGCTCAAG 683
 QY 222 GlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPhe 241
 Db 684 CGCTGGCGCGCGCTGCTGTCAGCGGTGGCGCGGTGTGCGAGGACCTGGCTGCTGCTGCT 743
 QY 242 LeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIle 261
 Db 744 ACCGCCCTGTCCCGCTGGCGCGCGGATCTCACCTACCGTGGCGCGCGCTGGCGCATC 803
 QY 262 ThrThrCysHisAspValHisAsn---ThrCysGluSerSerSerProPheGlnLeuTyr 280

Db 804 ATCACCTGCTTCGACGCTCTCAAGTGACGATGCTCCCCAGCGTGGCCATGTGGGCGGTG 863
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleTyrCys 300
Db 864 TTCTCTTCACCATCTTCATCCTGCTGTTCTCATCCGTTCTGATCACCCTGGCTTGT 923
Qy 301 TyrAlaAla-----IleIleArgThrLeuAsnAlaTyrAsp---HisArgTrp 315
Db 924 TACAGGCCACCATCTCAAGCTGTTCGCGACGAGGAGCGCACGCCCGGAGCAGCGG 983
Qy 316 LeuTrpTyrValIysAlaSerLeuLeuIleValIlePheThrIleCysPheAlaPro 335
Db 984 AGCGCCGCGTGGGCGCTGGCGGCTGTGTCGCTTGTTCACCTGCTTCGCCGCC 1043
Qy 336 SerAsnIleIleLeuIleIleHis---HisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 1044 AACAACTTCGTGCTCTGGCGCACATCGTGAGCGCGCTGTCTACGGCAAGAGC----- 1097
Qy 355 LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspPro 374
Db 1098 TACTACACGCTACAGCTCAGCCGTGTCTCAGCTGCCTCAACAACTGTCTGGACCCG 1157
Qy 375 PheLeuTyrPheLeuMetSerLys 382
Db 1158 TTGTTTATTACTTTGGCTCCCG 1181

Search completed: June 29, 2003, 08:19:28
Job time : 247.736 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 29, 2003, 07:29:54 ; Search time 2269.78 Seconds
(without alignments)
5231.316 Million cell updates/sec

Title: US-09-208-629F-6

Perfect score: 2157

Sequence: 1 CSMILQISXRLRDGTQVIM.....AYLTXXNDLRBQGQPSQRT 408

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09208629/runat_24062003.101625.12922/app.query.fasta_1.1166
-DB=GenEmbl -QFWT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09208629 @CGN 1 1 2827 @runat_24062003.101625.12922 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.ste.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.higo.hum.*
40: em.higo.mus.*
41: em.higo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	2142	99.3	1224	6	AR070426 Sequence
2	2142	99.3	1830	9	HSU92971 Human prote
3	1960.5	90.9	9273	9	AF374726 Homo sapi
4	1960.5	90.9	101887	9	AC026725 Homo sapi
5	1960.5	90.9	131370	9	AC026706 Homo sapi
6	1947	90.3	1102	6	AR070427 Sequence
7	1443.5	66.9	1224	6	AR070424 Sequence
8	1443.5	66.9	2409	10	MMU92972 Mus musculus
9	1424	66.0	1505	10	AF310076 Rattus no
10	1367.5	63.4	235097	2	AC110877 Mus muscu
11	959	44.5	1124	6	AR070425 Sequence
12	606.5	28.1	2713	10	WMPAR2MR Z48043 M.musculus
13	606.5	28.1	2732	6	AR012639 Sequence
14	606.5	28.1	2732	6	ARI17259 Sequence
15	606.5	28.1	2732	6	I87850 Sequence 60
16	605.5	28.1	1475	6	AR012637 Sequence
17	605.5	28.1	1475	6	ARI171257 Sequence
18	605.5	28.1	1475	6	I42454 Sequence 1
19	605.5	28.1	1475	6	I87848 Sequence 1
20	603.5	28.0	1477	10	BC025432 Mus muscu
21	603.5	28.0	1477	10	MMPROACR2 Z35158 M.musculus
22	603.5	28.0	199227	2	AC110876 Mus muscu
23	599.5	27.8	1680	5	XLJ09632 Xenopus lae
24	595	27.6	1428	10	RNU61373 U61373 Rattus norv
25	592.5	27.5	2813	9	BC012453 Homo sapi
26	592.5	27.5	88955	2	AC128225 Rattus no
27	592.5	27.5	151166	2	AC130634 Rattus no
28	591.5	27.4	2876	9	BC018130 Homo sapi
29	586.5	27.2	1451	9	HSU34038 Human prote
30	585.5	27.1	1124	9	HSU36753 Human prote
31	585.5	27.1	1255	6	AR012638 Sequence
32	585.5	27.1	1255	6	ARI171258 Sequence 3
33	585.5	27.1	1255	6	I42455 Sequence 3
34	585.5	27.1	1255	6	I87849 Sequence 3
35	585.5	27.1	1289	9	HSPAR2B Z49994 H.sapiens p
36	585.5	27.1	18351	9	AF400075 Homo sapi
37	585.5	27.1	52368	9	AC010621 Homo sapi
38	585.5	27.1	184536	2	AC068682 Homo sapi
39	585.5	27.1	204485	2	AC114962 Homo sapi
40	572.5	26.5	1414	6	AR012640 Sequence
41	572.5	26.5	1414	6	ARI171260 Sequence
42	572.5	26.5	1414	6	I87851 Sequence 62
43	555.5	25.8	153451	2	AC079576 Mus muscu
44	554.5	25.7	118914	2	AC079554 Mus muscu
45	543	25.2	1694	10	AF269246 Rattus no

ALIGNMENTS

RESULT 1

AR070426
 LOCUS AR070426 1224 bp DNA linear PAT 18-FEB-2000
 DEFINITION Sequence 4 from patent US 5892014.
 ACCESSION AR070426
 VERSION AR070426.1 GI:7221314
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE
 1 (bases 1 to 1224)
 Coughlin,S.R., Ishihara,H. and Connolly,A.
 TITLE DNA encoding a protease-activated receptor 3
 JOURNAL Patent: US 5892014-A 4 06-APR-1999;
 FEATURES
 Location/Qualifiers
 source 1..1224
 /organism="unknown"

BASE COUNT 303 a 316 c 221 g 384 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2,01e-165 Length: 1224
 Score: 2142.00 Matches: 405
 Percent Similarity: 99.26% Conservative: 0
 Best Local Similarity: 99.26% Mismatches: 3
 Query Match: 99.30% Indels: 0
 DB: 6 Gaps: 0

US-09-208-629F-6 (1-408) x AR070426 (1-1224)

Qy 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
 Db 1 TGCTCCATGATTACAGATTTCATAACGGTTAAGAGCGGACTCAGGTCACTAAATG 60
 Qy 21 LysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuProThrPheCysGlnSer 40
 Db 61 AAAGCCCTCATCTTTGCAGCTGCTGGCTCTCTGTTCTGTGGCCACTTTTGTCCAGAGT 120
 Qy 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
 Db 121 GGATGGAAATGATACAAACACTTGGCAAGCCCAACCTTACCCTTAAGACCTTTCGT 180
 Qy 61 GlyAlaProAsnSerPheGluPheProPheSerAlaLeuGluGlyTTPThrGly 80
 Db 181 GGAGTCCCCAAATCTTTTGAAGAGTTCCCTTTCTGCTTGAAGCTCGACAGGA 240
 Qy 81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
 Db 241 GCCACGATTACTGTAAATAATTAAGTGCCCTGAAGAAAGTCTTCACATCTCCATGTGAA 300
 Qy 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
 Db 301 AATGCTACCAATGGGTACCTGACAGCTCTTAAGTACTAACTGATACCTGCCACTCTAC 360
 Qy 121 LeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
 Db 361 CTCTGGTGTGTAGTTGGTGTGTCGGCCCAATGCTGTGACCTGTGGATGCTTTCTTC 420
 Qy 141 ArgThrArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
 Db 421 AGGACCAAGATCCATCTGTACCACTGTATTCTACACCAACCTGGGCCATTGCGATTTCCT 480
 Qy 161 PheCysValThrLeuProPheLeuIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
 Db 481 TTTTGTGTTACATGGCCCTTTAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTATT 540
 Qy 181 GlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
 Db 541 GGAGAGTCTGTGGCGGCCACACAGTCATCTTCTATGGCAACATGTAAGTCTCCATT 600
 Qy 201 LeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
 Db 601 CTGCTCTTGGCTGCATCAGCATCAACCGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660

Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
 Db 661 CGGGGCGCTGCCAAGCACACACCTATGCTTGGTAAACATGTGGACTGGTGGGCAACAGTT 720
 Qy 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrIleuValGlnProAsp 260
 Db 721 TTTCTATATATGCTGCCATTTTTCATCTACTGAAGCAGGAATATATCTTCTGAGCCAGAC 780
 Qy 261 IleThrThrCysHisAspValHisenThrCysGluSerSerSerProPheGlnLeuTyr 280
 Db 781 ATCACCACCTGCCATGATGTTTCCAAACACTTCGAGTCTCATCTCCCTTCCAACTCTAT 840
 Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
 Db 841 TACTTCATCTCTCTGGCATTCTTTGGATTCTTAATTCATTTGCTTATCATCTACTGCG 900
 Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
 Db 901 TATGAGCCATCATCGGACACTTAATGCATACGATCATAGATGGTTGTGGTATGTTAAG 960
 Qy 321 AlaSerLeuLeuIleLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
 Db 961 GCGAGTCTCTCATCTGTGATTTTACCATTGCTTTGCTCCAGCAATATATATCTT 1020
 Qy 341 IleIleHisHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
 Db 1021 ATATTCACCATGCTAACTACTACTACACACACTGATGGCTTATATTTATATATCTC 1080
 Qy 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
 Db 1081 ATAGCTTTGCTGCTGGTAGTCTTAATAGTGTGCTTAGATCCATCTCTTTATTTCTCATG 1140
 Qy 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
 Db 1141 TCAAAAACAGAAATCACTCCACTGCTTACCTTACAAAATAGTGAATGATCTTAGAGAA 1200
 Qy 401 GlnGlyGlnProSerGlnArgThr 408
 Db 1201 CAAGNACAGCCATCAGAGAAG 1224

RESULT 2
 HSU92971 1830 bp mRNA linear PRI 16-APR-1997
 LOCUS Human protease-activated receptor 3 (PAR3) mRNA, complete cds.
 DEFINITION U92971
 ACCESSION U92971.1 GI:1938374
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1830)
 Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
 Timmons,C., Tram,T. and Coughlin,S.R.
 TITLE Protease-activated receptor 3 is a second thrombin receptor in humans
 JOURNAL Nature 386 (6624), 502-506 (1997)
 MEDLINE 97242411
 PUBMED 9087410
 REFERENCE 2 (bases 1 to 1830)
 AUTHORS Ishihara,H., Connolly,A.J., Zeng,D., Kahn,M.L., Zheng,Y.W.,
 Timmons,C., Tram,T. and Coughlin,S.R.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAR-1997) CVRI, UCSF, 3rd and Parnassus, San Francisco, CA 94143, USA
 FEATURES
 Location/Qualifiers
 source 1..1830
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene 1..1830
 /gene="PAR3"
 CDS 145..1269
 /gene="PAR3"

/note="thrombin receptor; coagulation protease"

/codon_start=1

/product="protease-activated receptor 3"

/protein_id="AAC51218.1"

/db_xref="GI:1938375"

/translation="MKALIPAAAGILLLLPTFCQSGMENDTNLAKPTLPIKTFRGAP
PNSFEFGPSALSGWTGATITVKKCPESASHLVKNATMGYLTSLSKLIPALYL
LVFVGPVANATLWMLFFRTSICITVFYNLAIADFLFCVTLTPFKIAYHLGNWV
FGVLCRAITVIYGNMYSILLACISINRYLAI VHPFTYRGLPKHTVALVTCGLVW
ATVFLMLPFILKQYVILVQPDITTCVDVNTCESSPFOLYFISLAFPGFLIPFV
LIICYAAIIRTLNADHWLWVKASLLILVFTICFAPSNIILIIHHVYNNYND
GLYFVILALCIGLSNCLDPLFLYFLMSKTRNHSYALTK"

BASE COUNT 473 a 464 c 337 g 556 t

ORIGIN

Alignment Scores:

Pred. No.: 3 2e-165 Length: 1830
Score: 2142.00 Matches: 405
Percent Similarity: 99.26% Conservative: 0
Best Local Similarity: 99.26% Mismatches: 3
Query Match: 99.30% Indels: 0
DB: 9 Gaps: 0

US-09-208-629f-6 (1-408) x HSU92971 (1-1830)

Qy 1 CysSerMetIleLeuGlnIleSer***ArgLeuArgAspGlyThrGlnValIleLysMet 20
Db 88 TGCTCCATGATTACAGATTTTACAGTTTAAAGAGCGGGACTCAGGTCACTCAAAATG 147
Qy 21 LysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSer 40
Db 148 AAGAGCCCTCATCTTTGCAGCTGCTGCGCTCTCTCTCTGTTGCGCCACTTTTGTG 207
Qy 41 GlyMetGluAsnAspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPheArg 60
Db 208 GGCATGGAAATGATACAAACAACTTGGCAAGCCAACTTACCATTAAAGACTTTCGT 267
Qy 61 GlyAlaProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTyrThrGly 80
Db 268 GGAGCTCCCCAAATCTTTTGAAGAGTTCCCTTTCTGCTGGAGGCTGGACAGGA 327
Qy 81 AlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLys 100
Db 328 GCCACGATTACTGTAAATAATTAAGTCCCTGAAAGAGTCTTCACATCTCCATGTGAAA 387
Qy 101 AsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyr 120
Db 388 AATGCTACCAAGGGGTACCTGACAGCTCTTAAGTACTAACTGATACCTGCACTAC 447
Qy 121 LeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePhe 140
Db 448 CTCTGGTGTGTAGTTGGTGTCCCGCCCAATGCTGTGACCTGTGGATGCTTTTCTTC 507
Qy 141 ArgThrArgSerIleCysThrValPheTyrThrAsnLeuAlaIleAlaAspPheLeu 160
Db 508 AGGACCAAGATCCATCTGACCACTGTATTCTACACCAACCTGGCCATTGAGATTTCTT 567
Qy 161 PheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPhe 180
Db 568 TTTTGTGTTACATGGCTTTTAGATAGCTTATCATCTCAATGGAAACAACCTGGGTATT 627
Qy 181 GlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIle 200
Db 628 GGAGAGTCTGTGCGGGCCACACACAGTCATCTTCTATGGCAACATGACTGCTCCATT 687
Qy 201 LeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyr 220
Db 688 CTGCTCCTTCGCTGATCAGCATCAACCGTCACTGCGCCATCGTCCATCTTTCACCTAC 747
Qy 221 ArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrVal 240
Db 748 CGGGGCTGCCCAAGCACACTATGCTCTTGGTAAACATGTGGAGTGGTGGGCAACAGTT 807

Qy 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db 808 TTCTTATATATGCTGCCATTTTCTACTAGAACGAGGAATATATCTTCTTCCAGCAGAC 867
Qy 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGlnLeuTyr 280
Db 868 ATCACCACCTGTCATGATGTTTACAACTTGGAGTCTCATCTCCCTTCCAACTCTAT 927
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 928 TACTTCATCTCTTGGCATCTTTGGATTCTTAAATTCATTTGCTGCTTATCATCTACTGC 987
Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLys 320
Db 988 TATGAGCCATCATCCGACACTTAATGCATACGATCATAGATGGTGTGGTATGTTAAG 1047
Qy 321 AlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeu 340
Db 1048 GCGAGTCTCTCATCTTGTGATTTTACCAITTTGCTTCCCAAGCAATATATATCTT 1107
Qy 341 IleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeu 360
Db 1108 ATTATTCACCATGCTAACTACTTACAAACACACATGATGGCTTATATTTATATATCTC 1167
Qy 361 IleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMet 380
Db 1168 ATAGCTTTTGGCTGGGTAGTCTTAAATAGTTGCTTAGATCCATCTCTTTATTTCTCATG 1227
Qy 381 SerLysThrArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1228 TCAAAACCCAGAAATCACCTCCACTGCTTACCTTACAAAATAGTGAATGATCTTAGAGAA 1287
Qy 401 GlnGlyGlnProSerGlnArgThr 408
Db 1288 CAAGACACCCATCAGAGAACG 1311
RESULT 3
AF374726 9273 bp DNA linear PRI 01-JUN-2001
LOCUS Homo sapiens coagulation factor II receptor-like 2 (F2RL2) gene,
DEFINITION complete cds.
ACCESSION AF374726
VERSION AF374726.2 GI:14278710
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 9273)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,O. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REFERENCE 2 (bases 1 to 9273)
AUTHORS Rieder,M.J., Carrington,D.P., Chung,M.-W., Lee,K.L., Poel,C.L.,
Yi,O. and Nickerson,D.A.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2001) Molecular Biotechnology, University of
Washington, 1705 NE Pacific, Seattle, WA 98195, USA
REMARK Sequence update by submitter
COMMENT On Jun 1, 2001 this sequence version replaced gi:13991934.
To cite this work please use: SeattleSNPs. NHLBI program for
Genomic Applications. UW-FHCRC, Seattle, WA (URL:
http://pga.mbt.washington.edu).
FEATURES
source
1..9273
/organism="Homo sapiens"
/db_xref="taxon:9606"
repeat_region 120..221
/rpt_family="LIMB3"
/rpt_type=dispersed
variation 419

```

/ frequency="0.29"
/ replace="G"
483
variation
/ frequency="0.02"
/ replace="T"
506
variation
/ frequency="0.32"
/ replace="G"
507
variation
/ frequency="0.25"
/ replace="C"
512
variation
/ frequency="0.24"
/ replace="A"
535
variation
/ frequency="0.24"
/ replace="T"
937
variation
/ frequency="0.02"
/ replace="C"
969..7322
gene
/ gene="F2RL2"
join(969..1176,5701..7322)
mRNA
/ gene="F2RL2"
/ product="coagulation factor II receptor-like 2"
1007
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="G"
join(1113..1176,5701..6761)
CDS
/ gene="F2RL2"
/ notes="thrombin receptor-like 2"
/ codon_start=1
/ product="coagulation factor II receptor-like 2"
/ protein_id="AAK51564.1"
/ db_xref="GI:13991935"
/ translations="MKALIPAAAGLLLLPTFCQSGMENDTNNLAKPTLPKTFRGAP
PNSFERPPSALSGWTGATITVKIKCPESASHLVKNATGMVLTSSLTKLIPAYL
LVVVGVPANAVTLMWLPFRTSICITVFNIAIADFLCVTLPKIAYHLNGNW
FGVLKATVIFGNMYSILLACISINRYLAI VHPFTYRGLPKRTYALVTCGLW
ATVFLYMLPFILKQBYLVQPDITTCDDVHNTCESSPPQLYFISLAPFGLIPFV
LIICYAAIIRTNAYDHRWLVKASLLILVFTICFAPSNIILIIHANYYYNNTD
GLFYIYLIALLCLGSLNSCLDPFLYFLMSKTRHNSTAYLTK"
1156
variation
/ gene="F2RL2"
/ frequency="0.01"
/ replace="C"
1430
variation
/ gene="F2RL2"
/ frequency="0.06"
/ replace="G"
2055
variation
/ gene="F2RL2"
/ frequency="0.03"
/ replace="T"
2120
variation
/ gene="F2RL2"
/ frequency="0.01"
/ replace="T"
2136
variation
/ gene="F2RL2"
/ frequency="0.01"
/ replace="G"
2168
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="T"
2174
variation
/ gene="F2RL2"
/ frequency="0.01"
/ replace="C"
2250..2441
repeat_region
/ rpt_family="MER107"

/ rpt_type="dispersed"
2448
/ gene="F2RL2"
/ frequency="0.16"
/ replace="T"
2535
variation
/ gene="F2RL2"
/ frequency="0.38"
/ replace="T"
2800
variation
/ gene="F2RL2"
/ frequency="0.04"
/ replace="A"
2866
variation
/ gene="F2RL2"
/ frequency="0.35"
/ replace="T"
3116
variation
/ gene="F2RL2"
/ frequency="0.03"
/ replace="A"
3135
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="A"
3281
variation
/ gene="F2RL2"
/ frequency="0.04"
/ replace="A"
3386
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="A"
3543..4180
repeat_region
/ rpt_family="L2"
/ rpt_type="dispersed"
3553
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="T"
3565
variation
/ gene="F2RL2"
/ frequency="0.08"
/ replace="T"
3610
variation
/ gene="F2RL2"
/ frequency="0.01"
/ replace="C"
3694
variation
/ gene="F2RL2"
/ frequency="0.07"
/ replace="C"
3734
variation
/ gene="F2RL2"
/ frequency="0.32"
/ replace="T"
3850
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="A"
3976
variation
/ gene="F2RL2"
/ frequency="0.60"
/ replace="C"
4181..4526
repeat_region
/ rpt_family="MLT1D"
/ rpt_type="dispersed"
4223..4226
variation
/ gene="F2RL2"
/ frequency="0.02"
/ replace="C"
4264
variation
/ gene="F2RL2"

```

```

/frequency="0.32"
/replace="C"
4305
/genes="F2RL2"
/frequency="0.35"
/replace="G"
4464
/genes="F2RL2"
/frequency="0.01"
/replace="G"
4527...4827
/rpt_family="AluY"
/rpt_type=dispersed
4536
/genes="F2RL2"
/frequency="0.03"
/replace="A"
4580
/genes="F2RL2"

```

Alignment Scores:

```

Pred. No.: 1,2e-149 Length: 9273
Score: 1960.50 Matches: 375
Percent Similarity: 95.49% Conservative: 6
Best Local Similarity: 93.98% Mismatches: 13
Query Match: 90.89% Indels: 5
DB: 9 Gaps: 1
US-09-208-629f-6 (1-408) x AF374726 (1-9273)

QY 12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaGlyLeuLeu 31
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5620 AGAAATATTTCACAAATAATTAACAC-----ACTTCTCTTCTCTGAATATAC 5670

QY 32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAlaL 51
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5671 ATTATGTTAACTTTCTTCTCAATACAGCATGGAATAATGATACAACTTGCGCAA 5730

QY 51 ysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5731 AGCCAACTTACCAATTAAGACCTTCGTGGAGCTCCCCCAATCTTTTGAAGATTCC 5790

QY 71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5791 CCTTTCTGCTTGGAGGCTGGACAGGACCGCATTTACTGTAAAAAATTAAGTCCTG 5850

QY 91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerL 111
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5851 AAGAAAGTGCTTCACATCTCCATGTGAAATGCTACCATGGGGTACCTGACCATCTCT 5910

QY 111 euSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGlyValProAlaA 131
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5911 TAAGTAACTAAGTACCTGCCATCTACCTCTGTTGTTGTAGTTGGTGTCCCGCCA 5970

QY 131 snAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheT 151
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
5971 ATGCTGTGACCTGTGGATGCTTTCTTCAGGACAGATCCATCTGTACCATGTATTCT 6030

QY 151 yrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6031 ACACCAACCTGGGCATTCAGATTTCTTTTGTGTGTACATTTGCCCTTTAAGATAGCT 6090

QY 171 yrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysAtgAlaThrValI 191
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6091 ATCATCTCAATGGGAACCACTGGGTATTTGGAGAGGTCTGTGCGGGCCACCAAGTCA 6150

QY 191 lePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIleSerIleAsnArgT 211
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6151 TCTTCTATGCAACATGCTACTGCTCCATCTGCTCTGCTGCTGCTGCTGCTGCTGCTG 6210

QY 211 yrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuV 231
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6211 ACCTGGCCATCGTCCATCTCTTACCTTACCTGCGGGGCTGCGGCAAGCACACCTATGCTTGG 6270

```

```

QY 231 alThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuL 251
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6271 TAACATGTGCTGGTGTGGCAACAGTTTCTTATATATGCTGCCATTTTTCATCTGA 6330

QY 251 ysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrC 271
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6331 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCACACACTT 6390

QY 271 ysGluSerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheL 291
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6391 GCGAGTCTCATCTCCCTTCCAACTTATATCTTCTCTCTGCGCATTTTGGATTCT 6450

QY 291 euIleProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaT 311
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6451 TAATTCATTTGTGCTTATCATCTACTGCTATGAGCCATCATCCGACACTTAAATGCAT 6510

QY 311 yrAspHisArgTrpLeuTyrValLysAlaSerLeuLeuLeuValIlePheThrI 331
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6511 ACGATCATAGATGGTGTGTATGTTAAGCGAGTCTCCTCATCTCTGTGATTTTACCA 6570

QY 331 leCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyAsnA 351
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6571 TTTGCTTGTCTCAAGCAATATATTCTTATTATTCACCATGCTAACTACTACTACAACA 6630

QY 351 snThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 371
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6631 ACACGTAGTGGCTTATATTATATATATCTCATAGCTTGTGCTGGGTAGTCTTAATAGTT 6690

QY 371 ysLeuAspProPheLeuTyrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrL 391
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6691 GCTTAGATCATCTCTTTTATTTCTCATGTCAAAACAGAAATCACTCCACTGCTTACC 6750

QY 391 euThrLys*****AsnAspLeuAtGluGlnGlyGlnProSerGlnArgThr 408
Db |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6751 TTACAAATATAGTGAATGATCTTAGAGAAACAGGACGACCATCACAGAACG 6803

RESULT 4
AC026725/c
LOCUS AC026725 101887 bp DNA linear PRI 21-JUL-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2236F14, complete sequence.
ACCESSION AC026725
VERSION AC026725.7 GI:14993696
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 101887)
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS 3 (bases 1 to 101887)
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE
AUTHORS 4 (bases 1 to 101887)
TITLE Direct Submission
JOURNAL Submitted (21-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT
On Jul 21, 2001 this sequence version replaced gi:14572130.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;

```

Estimated Total Number of Errors is 0.1.

STS Content:
SHGC-102125 G56565
SHGC-101620 G56147
SHGC-83998 G53567.

FEATURES

Location/Qualifiers
1. .101887
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2236F14"
BASE COUNT 29200 a 20058 c 21105 g 31524 t
ORIGIN

Alignment Scores:

Pred. No.: 1.88e-148 Length: 101887
Score: 1960.50 Matches: 375
Percent Similarity: 95.49% Conservative: 6
Best Local Similarity: 93.98% Mismatches: 13
Query Match: 90.89% Indels: 5
DB: 9 Gaps: 1

US-09-208-629F-6 (1-408) x AC026725 (1-101887)

Qy 12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeu 31
Db 93557 AGAAATTATTACAAATAATAAACAC-----ACTTCTCTTCTCTTGAACATATAC 93507
Qy 32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAsnLeuAlaL 51
Db 93506 ATTTATGTTAATACTTCTTCTTCAATTACAGGCATGGAAATGATACAAACAACTTGCCAA 93447
Qy 51 ysrProThrIeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
Db 93446 AGCAACCTTACCATTAAGACCTTTGCTGGAGCTCCCAATCTCTTTGAAGAGTTCC 93387
Qy 71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
Db 93386 CTTTTCTGCTTGGAGGCTGGACAGGACCCAGGATTACTGTAAAAATTAAGTGGCCCTG 93327
Qy 91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrrLeuThrSerSerL 111
Db 93326 AAGAAAGTGTTCACATCTCCATCTGAAAAATGCTACCATGGGGTACCTGACCAGCTCCT 93267
Qy 111 euSerThrLysLeuIleProAlaIleTyrrLeuLeuValPheValGlyValProAlaA 131
Db 93266 TAAGTACTAACTGATACCTGCCATCTACCTCTCTGGTGTGTAGTGGTGTCCCGGCCA 93207
Qy 131 snAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheT 151
Db 93206 ATGCTGTGACCTGTGGATGCTTTCTTCAGGACCATCATCTGTACCATGTATTCT 93147
Qy 151 yrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
Db 93146 ACACCAACCTGGCCATTCAGATTTCTTTTGTGTGTACATTCCTCCCTTTAAGATAGCTT 93087
Qy 171 yrHisLeuAnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrThrValI 191
Db 93086 ATCATCTCAATGGGAACAACATGGGTATTTGGAGAGGTCCTGTGCGGGGCCACACAGTCA 93027
Qy 191 lePheTyrrGlyAsnMetTyrrCysSerIleLeuLeuAlaCysIleSerIleAsnArgT 211
Db 93026 TCTTCTATGGCAACATGATCTGCTCCATTTCTGCTCTGCTGCTGATCAGCATCAACCGCT 92967
Qy 211 yrLeuAlaIleValHisProPheThrTyrrArgGlyLeuProLysHisThrTyrrAlaLeuV 231
Db 92966 ACCTGGCCATCGTCCATCTTCACTTACCTTACCTGGGGCTCGCCAAAGCACCATATGCTTGG 92907
Qy 231 alThrCysGlyLeuValTrpAlaThrValPheIleTyrrMetLeuProPhePheIleLeuL 251
Db 92906 TAACATGGAGCTGGTGGGCAACAGTTTCTTATATATGCTGCTGCTGCTGCTGCTGCTGCTG 92847
Qy 251 ysGlnGluTyrrTyrrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrC 271

Db 92846 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACCTGCCATGATGTTCAACACATT 92787
Qy 271 ysGluSerSerSerProPheGlnLeuTyrrPheIleSerLeuAlaPhePheGlyPheL 291
Db 92786 GCGAGTCTCTCATCTCCCTTCCAACTCTATTACTTCTATCTCTTGGCATTTTGGATTCT 92727
Qy 291 euIleProPheValLeuIleIleTyrrCysTyrrAlaAlaIleIleArgThrLeuAsnAlaT 311
Db 92726 TAAATCCATTGTGCTTATCATCTACTGCTATGCGACCATCATCCGACACTTAATGCAT 92667
Qy 311 yrAspHisArgTrpLeuTrpValLysAlaSerLeuLeuIleValIlePheThrI 331
Db 92666 ACGATCATAGATGGTGTGTATGTAAGCGAGTCTCTCATCTCTGTGATTTTACCA 92607
Qy 331 leCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrrTyrrAsnA 351
Db 92606 TTTGCTTTGCTCCAGCAATATTATTCTTATTATTCACCATGCTAACTACTACTACAACA 92547
Qy 351 snThrAspGlyLeuTyrrPheIleTyrrLeuIleAlaLeuCysLeuGlySerLeuAsnSerC 371
Db 92546 ACATGATGGCTTATATTATATATCTCATAGCTTGTGCTGGTAGTCTTAATAGTT 92487
Qy 371 ysLeuAspProPheLeuTyrrPheLeuMetSerLysThrArgAsnHisSerThrAlaTyrl 391
Db 92486 GCTTAGATCCATCTCTTATTCTTCTCATGTCAAAACACAGAAATCACTCCACTGCTTACC 92427
Qy 391 euThrLys*****AsnAspLeuArgGluGlnGlyGlnProSerGlnArgThr 408
Db 92426 TTACAAATAGTGAATGATCTTAGAGAACAAAGGACAGCCATCACAGAGAACG 92374

RESULT 5

AC026706 131370 bp DNA linear PRI 21-JUN-2001
LOCUS Homo sapiens chromosome 5 clone CTD-2068N7, complete sequence.
DEFINITION AC026706
ACCESSION AC026706.5 GI:14518404
VERSION HTG.
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 131370)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 131370)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 131370)
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT On Jun 21, 2001 this sequence version replaced gi:13677033.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >40 99.6% of Sequence;
Estimated Total Number of Errors is 0.3.
SIS Content:
WI-11362 G24506.
Location/Qualifiers
1. .131370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2068N7"
BASE COUNT 40462 a 26452 c 25592 g 38864 t
ORIGIN

FEATURES

source
1. .131370
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2068N7"
BASE COUNT 40462 a 26452 c 25592 g 38864 t
ORIGIN

Alignment Scores:
Pred. No.: 2,52e-148 Length: 131370
Score: 1960.50 Matches: 375
Percent Similarity: 95.49% Conservative: 6
Best Local Similarity: 93.98% Mismatches: 13
Query Match: 90.89% Indels: 5
DB: 9 Gaps: 1

US-09-208-629F-6 (1-408) x AC026706 (1-131370)

QY 12 ArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeu 31
Db 60121 AGAAATATTACAAATATAAACAC-----ACTTCTTTTCTCTTGAACATATAC 60171

QY 32 LeuLeuLeuProThrPheCysGln--SerGlyMetGluAsnAspThrAsnAsnLeuAlaL 51
Db 60172 ATTTATGTTAATACTTCTTTTCAATTACAGGCATGGAATAATGATACAAACAACCTGGCAA 60231

QY 51 YsProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheP 71
Db 60232 AGCAACCTTACCATTAAAGACCTTCGTGGAGCTCCCAAAATCTTTTGAAGAGTCTC 60291

QY 71 roPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProG 91
Db 60292 CCTTTCGTCTGGAAAGCTGGACAGGACACGATTACTGTAAAAATTAAGTGCCTG 60351

QY 91 luGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyLeuThrSerSerL 111
Db 60352 AAGAAAGTGTCTCACATCTCCATGTGAAAAATGCTACCAATGGGTACCTGACGAGTCT 60411

QY 111 euSerThrLysLeuIleProAlaIleTyLeuLeuValPheValValGlyValProAla 131
Db 60412 TAAGTACTAAACTGATACCTGCCATCTACCTCTCTGGTGTGTGGTGTCCCGCCA 60471

QY 131 snAlaValThrLeuTrpMetLeuPheArgThrArgSerIleCysThrThrValPheT 151
Db 60472 ATGCTGTGACCTGTGGATCTTTCTTCAGACACGATCCATCTGTACCACTGTATTCT 60531

QY 151 YrThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaT 171
Db 60532 ACACCAACCTGGCCATTGACAGATTTCTTTTGTGTGTATCATTTGCCCTTTTAAGATAGCTT 60591

QY 171 YrHisLeuAnGlyAsnAsnTrpValPheGlyGluValLeuCysArgAlaThrValI 191
Db 60592 ATCATCTCAATGGAAACACTGGGTATTTGGAGAGGTCCTGTGCCGGCCACACAGTCA 60651

QY 191 lePheTyrglyAsnMetTyrcysSerileLeuLeuAlaCysIleSerIleAsnArgT 211
Db 60652 TCTTCTATGCAACATGATCTGCTCCATCTGCTCTTGGCTGCATCAGCATCAACCGCT 60711

QY 211 YrLeuAlaIleValHisProPheThrTyrglyLeuProLysHisThrTyrrAlaLeuV 231
Db 60712 ACCTGGCCATCGTCCATCTCTTCACTACCGGGGCTGCCAAGCACACCTATGCTCTGG 60771

QY 231 alThrCysGlyLeuValTrpAlaThrValPheLeuTyrrMetLeuProPhePheIleLeuL 251
Db 60772 TAACATGTGACTGGTGTGGCAACAGTTTCTTATATATGCTGCCATTTTTCATACTGA 60831

QY 251 YsGlnGluTyrrTyrrLeuValGlnProAspIleThrTyrcysHisAspValHisAsnThrC 271
Db 60832 AGCAGGAATATTATCTTGTTCAGCCAGACATCACCACTGCCATGATGTTCACCAACACT 60891

QY 271 YsGluSerSerPropheGlnLeuTyrrThrPheIleSerLeuAlaPhePheGlyPheI 291
Db 60892 GCGAGTCTCATCTCCCTTCCAACTATTACTTTCATCTCTTGGCATCTTTTGGATTCT 60951

QY 291 euileProPheValLeuIleIleTyrrCysTyrrAlaAlaIleIleArgThrLeuAsnAlaT 311
Db 60952 TAATTCATTGTGCTTATCATCTACTGTCTATGACCCATCATCCGGACACTTAATGCA 61011

QY 311 YrAspHisArgTrpLeuTrpTyrrValLysAlaSerLeuLeuIleLeuValIlePheThrI 331

Db 361 CTTTGTGTTTACATGCTTAAAGATAGCTTATCATCTCAATGGGAACAACCTGGGTA 420
QY 180 PheGlyGluValLeuCysArgAlaThrValIlePheThrCysSer 199
Db 421 TTTGGAGAGGTCTGTGGCGGGCCACACAGTCTCTTATGGCAACATGTACTCTCC 480
QY 200 IleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThr 219
Db 481 ATTCTGCTCTTGGCTGCATCAGCANTCAACCGTACTCTGGCCATCTGTCATCTTCACC 540
QY 220 TyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValThrAlaThr 239
Db 541 TACCGGGCTGCCCAAGCACACCTATGCTTGGTAACATGTGAGCTGGTGGGCAACA 600
QY 240 ValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnPro 259
Db 601 GTTCTTCTATATATGCTGCCATTTTTCATCTGAAGCAGGAATATATCTTGTTCAGCCA 660
QY 260 AspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeu 279
Db 661 GACATCACCACTGCGCATGTATGTTCAACACTTGGAGTCTCATCTCCCTTCCACTC 720
QY 280 TyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleTyr 299
Db 721 TATTACTTCTCTCTGTCATCTTGGATCTTAAATCCATTTGTGCTTATCATCTAC 780
QY 300 CysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTyrVal 319
Db 781 TGTATGACGCCATCATCCGGACACTTAATGCATACGATCATAGATGGTGTGGTATGT 840
QY 320 LysAlaSerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIle 339
Db 841 AAGCGAGTCTCTCTCATCTTGTGATTTTACATTTGCTTCTCCAGCATATATT 900
QY 340 LeuIleIleHisAlaAsnTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyr 359
Db 901 CTTATTATTACCATGTACTACTACTACCAACACACTGATGGCTATATTTATATAT 960
QY 360 LeuIleAlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeu 379
Db 961 CTATAGCTTTGGCTGGGTAGTCTTAATAGTTGCTTAGATCCATCTCTTTATTTTCTC 1020
QY 380 MetSerLysThrArgAsnHisSerThrAlaTyrLeuThrLys****AsnAspLeuArg 399
Db 1021 ATGTCAAAACCAAGAAATCACTCACTGCTTACTTACAAATAGTGAATGATCTAGA 1080
QY 400 GluGlnGlyGlnProSerGln 406
Db 1081 GAACAAGGACAGCCATCACAG 1101

RESULT 7
LOCUS AR070424 1224 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 1 from patent US 5892014.
ACCESSION AR070424
VERSION AR070424.1 GI:7221312
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1224)
AUTHORS Coughlin, S.R., Ishihara, H. and Connolly, A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 1 06-APR-1999;
FEATURES Location/Qualifiers
source 1.1224
BASE COUNT 299 a 341 c 239 g 345 t
ORIGIN
Alignment Scores:
Pred. No.: 1.33e-108 Length: 1224

Score: 1443.50 Matches: 264
Percent Similarity: 83.74% Conservative: 45
Best Local Similarity: 71.54% Mismatches: 59
Query Match: 66.92% Indels: 1
DB: Gaps: 1
US-09-208-629F-6 (1-408) x AR070424 (1-1224)
QY 14 GlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeu 33
Db 33 GGGTCTCAGGACATCAAGATGAAATCCTTATCTTGGTTCAGCTGGGTGCTGTTCTG 92
QY 34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAlaLysProThr 53
Db 93 CCAGTCATCTGTTTGCNAAGTGCATA---AATGTTTCAGACAACCTCAGCAAGCCACC 149
QY 54 LeuProIleLysThrPheArgGlyAlaProProAsnSerPheGluGluPheProPheSer 73
Db 150 TTAACCTATTAGAGTATTAATGGGGTCCCAAAATACCTTTGAAGAAATTCACACTTCT 209
QY 74 AlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93
Db 210 GACATAGAGGCTGGACAGGACCCACCACTATAAAAGCGAGTGTCCGAGGACAGT 269
QY 94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThr 113
Db 270 ATTTCAACTCTCCACGTGAATAATGCTACCATAGGATACCTGAGAAGTTTCTTAAGTACC 329
QY 114 LysLeuIleProAlaIleTyrLeuLeuValPheValGlyValProAlaAlaVal 133
Db 330 CAAGTGATACCTGCCATCTATCTCTGTTTGTGGTGGTGTACCATCCCAACATCGTG 389
QY 134 ThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThrThrValPheTyrThrAsn 153
Db 390 ACCCTGTGGAACACTCTCTTAAGGACCAATCCATCAGTCTGCTCATCTTTCACCAAC 449
QY 154 LeuAlaIleAlaAspPheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeu 173
Db 450 CTGGCCATCGCAGATCTCTCTTCTGTGTGCACACTGCCATTTAAGATCGCTACCATCTC 509
QY 174 AsnGlyAsnAsnTrpValPheGlyGluValLeuLysArgAlaThrThrValIlePheTyr 193
Db 510 AATGGCAACAACATGGGTATTTGGCGAGGTCAATGGCGGATCACCACGGTCTGTTTCTAC 569
QY 194 GlyAsnMetTyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAla 213
Db 570 GGCAACATCTACTGCGCTATCTCTCATCTTGCATGGGATCAACCGCTACCTACCTGGCC 629
QY 214 IleValHisProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCys 233
Db 630 ACGGCTCACCTTTACATACACAGAGCTGCCAAACGCGAGCTTCTCTTGTCTCATGTGT 689
QY 234 GlyLeuValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLysGlnGlu 253
Db 690 GGCATAGTGTGGGTCAATGTTTCTTATACATGCTGCCCTTTTGTCTCATCTCTGAAGCAGG 749
QY 254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
Db 750 TACCACCTCGTCACTCAGAGATCACCATCGCAGATGTCTGTCGACGGTGGAGTCC 809
QY 274 SerSerProPheGlnLeuTyrTyrPheIleSerLeuAlaPhePheGlyPheLeuIlePro 293
Db 810 CCATCATCTCTCCGATTCCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCCG 869
QY 294 PheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAlaTyrAspHis 313
Db 870 TTTGTGATCATCTCTCTGTTTACAGACTCTCATCCACAACTTAAATCAAGAGTCCG 929
QY 314 ArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPhe 333
Db 930 ATATGGCTGGGTACATCAAGGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTT 989
QY 334 AlaProSerAsnIleIleLeuIleIleHisHisAlaAsnTyrTyrTyrAsnAsnThrAsp 353

RESULT 9	AF310076	1505 bp	mRNA	linear	ROD 02-NOV-2001
LOCUS	AF310076				
DEFINITION	Rattus norvegicus protease activated receptor 3 mRNA, complete cds.				
ACCESSION	AF310076				
VERSION	AF310076.1	GI:16588398			
KEYWORDS					
SOURCE	Rattus norvegicus.				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
REFERENCE	1 (bases 1 to 1505)				
AUTHORS	Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.				
TITLE	Cloning of the rat protease activated receptor isoforms 3 and 4				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1505)				
AUTHORS	Chien,E.K., Marietti,S., Mendoza,J. and Phillippe,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-OCT-2000) Obstetrics and Gynecology, University of Chicago, 5841 South Maryland Ave, MC2050, Chicago, IL 60637, USA				
FEATURES	Location/Qualifiers				
source	1..1505				
	/organism="Rattus norvegicus"				
	/strain="Sprague-Dawley"				
	/db_xref="taxon:10116"				
	/tissue_type="spleen"				
	63..1169				
	/note="PAR3; G-protein coupled receptor"				
	/codon_start=1				
	/product="protease activated receptor 3"				
	/protein_id="AAU26789.1"				
	/db_xref="GI:16588399"				
	/translations="MEMKVLILVGRLFLPTVCOSGMKHVSDNSALTAESFNGNEH SFVEPLDSLEGWGTATTTIKACPEESITLHVNNATMGYLRSSLSITKVIPIAIIYL FVIGPANIVTLKLSRTKSLCLVIFHTNLADLLFCVTLPEKAYHLNGNDWVFG EVMCRVTPAYGMYCAIILLTCMGINRYLATVHPPTVKLPKRFTLLMCGVVMV VLYMLPLALKOBYHQGFTICHDVHDTCESPLFPFYFVSIAFFGLFPFVVS VLYTLHLKNAQDRKWLRYIKAVLLILVIFTICFAPTNIIILIIHHANNYYISNTDSL YFMYLALCLGSLNSCLDPFLYFIMSKIVDQLTS"				
CDS	BASE COUNT 359 a 394 c 316 g 436 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	6.48e-107	Length:	1505		
Score:	1424.00	Matches:	266		
Percent Similarity:	80.05%	Conservative:	73		
Best Local Similarity:	68.91%	Mismatches:	43		
Query Match:	66.02%	Indels:	4		
DB:	10	Gaps:	2		
US-09-208-629F-6 (1-408) x AF310076 (1-1505)					
Qy	14 GlyThrGlnValIleLysMetLysAlaLeuPheAlaAlaGlyLeuLeuLeu 33				
Db	51 GGGTCTCAACATGGAGATGAAGTCTTATCTGTGGTTGGGTAGCTGCTGTTCTG 110				
Qy	34 LeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsnLeuAlaLysProThr 53				
Db	111 CCGACCACTGTTTGCCAAAGTGGCATGAACATGTTTCAGACAACTCAGCC----- 161				
Qy	54 LeuProLysThrPheArgGlyAlaProAsnSerPheGluGluPheProPheSer 73				
Db	162 TTAACCTGCTGAGAGTTTAAATGGC--AAGCAACATTCCTTTGAAGAAATTCCTTCT 218				
Qy	74 AlaLeuGluGlyTyrThrGlyAlaThrIleThrValLysIleLysCysProGluGluSer 93				
Db	219 GACATAGGGCTGGACAGGAGCCACCAACTATAAAGCGAAGTGTCCGAGAAAGC 278				
Qy	94 AlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeuThrSerLeuSerThr 113				
Db	279 ATTCAACTCTCATGTGAATAATGCTACCATGGGATACCTGGAAGTCTCTTAAGTACC 338				

REFERENCE 1 (bases 1 to 235097)
AUTHORS Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and Cheng, J.F.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 235097)
AUTHORS Martin, J., Hosseini, R., Peng, Y., Peng, Z., Rubin, E.M. and
Cheng, J.-F.F.
TITLE Direct Submission
JOURNAL Unpublished
COMMENT Draft Sequence Produced by Berkeley PGA
PGA.lbl.gov
Center Code: PGABERK.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
* 1 2480: contig of 2480 bp in length
* 2481 2580: gap of unknown length
* 2581 100301: contig of 97721 bp in length
* 100302 100401: gap of unknown length
* 100402 235097: contig of 134696 bp in length.
FEATURES
source
1..235097
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-345121"
BASE COUNT 64117 a 51797 c 51868 g 67115 t 200 others
ORIGIN
Alignment Scores:
Pred. No.: 8,47e-100 Length: 235097
Score: 1367.50 Matches: 252
Percent Similarity: 82.19% Conservative: 48
Best Local Similarity: 69.04% Mismatches: 65
Query Match: 63.40% Indels: 1
DB: 2 Gaps: 0
US-09-208-629F-6 (1-408) x AC110877 (1-235097)
QY 44 AsnAspThrAsnAsnLeuAlaLysProThrLeuProLeuLysThrPheArgGlyAlaPro 63
Db 228794 AATGTTTCAGACAACTCAGCAAGCCCACTTAACTATTAAAGAGTTTAAATGGGGTCCC 228735
QY 64 ProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIle 83
Db 228734 CAAATACCTTTGAAGAAATCCCACTTTCTGACATAGAGGGCTGGACAGGACCAACACA 228675
QY 84 ThrValLysLeuLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThr 103
Db 228674 ACTATAAAGCGGAGTGTCGCGAGGACAGTATTTCACCTCCACCGTGAATAATGTACC 228615
QY 104 MetGlyTrpLeuThrSerSerLeuSerLeuThrLysLeuLeuProAlaLeuVal 123
Db 228614 ATAGGATACCTGAGAGTTCCTTAAGTACCACCACTGATGATCTCCATCTACATCTGCTG 228555
QY 124 PheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArg. 143
Db 228554 TTGTGGTTGGTGTAACAGCAACATCGTCACCTGTGGAACTCTCTTAAGGACCAAA 228495
QY 144 SerLeuCysThrThrValPheThrAsnLeuAlaLeuAlaAspPheLeuPheCysVal 163
Db 228494 TCCATCAGTCTGTGCTCTTTTCACCACTGGCCATCGCAGATCTCTTTTCTGTGTC 228435
QY 164 ThrLeuProPheLysLeuAlaThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 193
Db 228434 ACATGCCATTTAAGATCGCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 228375

QY 184 LeuCysArgAlaThrThrValIlePheThrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
Db 228374 ACCTGCCGATCACCACGGTTGTTTTCTACGGCAACATGTACTCGCTATCTCTGATCCTC 228315
QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db 228314 ACCTGCATGGCATCAACCGCTACCTGGCCACGGCTCACCCCTTTCATACACCAAGAGCTG 228255
QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyr 243
Db 228254 CCCAAGCAGCTTCTCCATGCTCATGTGGCATGTGGGTGATGTTTCTTATATAC 228195
QY 244 MetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThr 263
Db 228194 ATGCTCCCTTTGTCATCTCCTGAAGCAGGAGTACCACTCGTCCACTCCGAGATCACACC 228135
QY 264 CysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyrTyrPheIle 283
Db 228134 TGGCAGATGTGTTGACGGTGGAGTCCCATCATCTCTCCGATCTACTACTTCTGTC 228075
QY 284 SerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCysTyrAlaAla 303
Db 228074 TCCTTAGCATTTCTTGGTTCCTCATCCGTTTGTGATCATCATCTCTGTTACACGACT 228015
QY 304 IleIleArgThrLeuAsnAlaTyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeu 323
Db 228014 CTCATCCAAACTTAATCAAAAGGATCGGATGCTGGCTTACATCAAGGCGCTCCTC 227955
QY 324 LeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleIleIleHis 343
Db 227954 CTCATCCTTGTGATTTCACAATTTGCTTTGCCCCCAACATCATCTCTGTAATCCAC 227895
QY 344 HisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeu 363
Db 227894 CATGCCAATCTACTTACCACATACCGACAGCTTGTATGTATCTTATGTCTCTG 227835
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 227834 TGCCTGGGAGCTGAATAGCTGCCTAGATCCATCTTACTTCTGATGTCGAAAGTT 227775
QY 384 ArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGluGlnGlyGln 403
Db 227774 GTAGATC--AGCTTAATCTTGTAGTACGAATGCAAGACCACTTAGACCAAGGAGAG 227717
QY 404 ProSerGlnArgThr 408
Db 227716 ATATCTGGGAAGAG 227702
RESULT 11
AR070425
LOCUS 1124 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 2 from patent US 5892014.
ACCESSION AR070425
VERSION AR070425.1 GI:7221313
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1124)
AUTHORS Coughlin, S.R., Ishihara, H. and Connolly, A.
TITLE DNA encoding a protease-activated receptor 3
JOURNAL Patent: US 5892014-A 2 06-APR-1999;
FEATURES
source
1..1124
/organism="unknown"
BASE COUNT 251 a 314 c 210 g 326 t 23 others
ORIGIN
Alignment Scores:
Pred. No.: 3.09e-69 Length: 1124
Score: 959.00 Matches: 185
Percent Similarity: 73.99% Conservative: 34
Best Local Similarity: 62.50% Mismatches: 53

[illegible]

```
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db 400 TTTTCCATCGATGAGTTC-----417
QY 85 ValysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db 418 -----TCTGCGTCC-----426
QY 105 GlyTyrLeuThrSerSerLeuSerThrLysLeuLeuPheProAlaIleTyrLeuLeuValPhe 124
Db 427 ---ATCCTCACCGGAGCTGACACGGTCTTTCTCCGGTCTGTACATATTATGTGTTT 483
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSer 144
Db 484 GTGATTGGTTGGCCAGTAAGCATGGCCCTCTGGATCTTCTTCCGAACGAAGAAG 543
QY 145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db 544 AAACACCCCGCGTGAATTATACATGGCAACCTGGCCCTTGGCCGACCTCTCTCTGTGCATC 603
QY 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db 604 TGGTTCCCTCCCTGAAGATCTCTACACCTACATGGCAACAACTGGGTCTACGGGAGGCC 663
QY 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeu 203
Db 664 CTGTGCAGGTGCTCATTTGCTTCTATGTGTAAACATGATTTGCTTCCATCTCTTCATG 723
QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db 724 ACCTGGCTCAGGTGTCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
QY 224 ProlYsHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
Db 775 CCCAGGAAGAAG---GCAACATACCGCGTTGGCGTCTCTTGGCAATCTGGCTCTCGATT 831
QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
Db 832 TTTCTGTGCACCATCCCTTTGTATGTCATGAAGACAGACCATCTACATTCAGCATTTGAAC 891
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
Db 892 ATCACCACTGTACGATGTG-----CTGCCTGAGGAGGTATTG 930
QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
Db 931 GTGGGGGACATGTTCAATTACTTCTCTACCTGCGCATTTGAGTCTTCTCTGTTTCCCGGCC 990
QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleArgThrLeuAsnAla-----310
Db 991 CTCCTTACTGATCTGCTACGCTCATGATCAAGACGCTCGCTCTTCTGCTATGGAT 1050
QY 311 -----TyrAspHisArgTyrLeuTyrValLysAlaSerLeuLeuIleValIle 328
Db 1051 GAACACTCAGAGAACAAGGAGGAGGCTATCCGACTCATCATCCGCTGCTGGCCATG 1110
QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
Db 1111 TACTTCTATCTGCTTTGCTCTAGCAACCTTCTGCTAGTGTATTTCTTCTTAATCAAA 1170
QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
Db 1171 ACCCAGAGGAGGAGCAGCTACGCTCTACCTTGTGCGCTCTGCTGCTGCGCCCTC 1230
QY 369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
Db 1231 AACAGTGCATAGACCCCTTTGTTCTATTACTTTGTTCTCAAAAGATTTCAGGATCAGCC 1290
RESULT 17
ARI171257 1475 bp DNA linear PAT 17-DEC-2001
LOCUS Sequence 1 from patent US 6297026.
```

```
ACCESSION ARI171257
VERSION ARI171257.1 GI:17910207
KEYWORDS Unknown.
SOURCE ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin,J. and Scarborough,R.M.
TITLE Nucleic acids encoding the C140 receptor
JOURNAL Patent: US 6297026-A 1 02-OCT-2001;
FEATURES Location/Qualifiers
source 1..1475
/organism="unknown"
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN
Alignment Scores:
Pred. No.: 2,41e-40 Length: 1475
Score: 605.50 Matches: 140
Percent Similarity: 52.50% Conservative: 70
Best Local Similarity: 35.00% Mismatches: 117
Query Match: 28.07% Indels: 74
DB: 6 Gaps: 13
US-09-208-629F-6 (1-408) x ARI171257 (1-1475)
QY 18 IleLysMetLysAlaLeuIlePheAlaAlaGlyLeuLeuLeuLeuProThr---36
Db 221 GTTCATCTGAAATGTTCCATTTAAACACAGACAGCGCTTACTGTGGACCATTTATCTCAG 280
QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla-----50
Db 281 TTAATGATCTGCTCGCTTCTTTGTACAGACGCAACACAGTAAGGAAGAGTCTTA 340
QY 51 -----LysProThrLeuProfile-----LysThrPheArgGlyAlaProPro 64
Db 341 TTGGCAGATTAGAAACCCAGCCTCCAATCACTGGGAAAGGGTTCGGTA-GAACCGAGC 399
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThr 84
Db 400 TTTTCCATCGATGAGTTC-----417
QY 85 ValLysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db 418 -----TCTGCGTCC-----426
QY 105 GlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPhe 124
Db 427 ---ATCCTCACCGGAGGAGCTGACACGGTCTTTCTTCCGGTCTGTACATATTATGTGTTT 483
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSer 144
Db 484 GTGATTGGTTGGCCAGTAAGCATGGCCCTCTGGATCTTCTTCCGAACGAAGAAG 543
QY 145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
Db 544 AAACACCCCGCGTGAATTATACATGGCAACCTGGCCCTTGGCCGACCTCTCTCTGTGCATC 603
QY 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db 604 TGGTTCCCTCCCTGAAGATCTCTACACCTACATGGCAACAACTGGGTCTACGGGAGGCC 663
QY 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeu 203
Db 664 CTGTGCAGGTGCTCATTTGCTTCTATGTGTAAACATGATTTGCTTCCATCTCTTCATG 723
QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
Db 724 ACCTGGCTCAGGTGTCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
QY 224 ProlYsHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
Db 775 CCCAGGAAGAAG---GCAACATACCGCGTTGGCGTCTCTTGGCAATCTGGCTCTCGATT 831
```



```

DEFINITION Sequence 1 from patent US 5716789.
ACCESSION I87848
VERSION I87848.1 GI:3407788
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1475)
AUTHORS Sundelin, J. and Scarborough, R.M.
TITLE Method to determine ligands, agonist and antagonist of C140
receptor
JOURNAL US 5716789-A 1 10-FEB-1998;
FEATURES Location/Qualifiers
source 1. 1475
BASE COUNT 341 a 414 c 319 g 401 t
ORIGIN
Alignment Scores:
Pred. No.: 2,41e-40 Length: 1475
Score: 605.50 Matches: 140
Percent Similarity: 52.50% Conservative: 70
Best Local Similarity: 35.00% Mismatches: 117
Query Match: 28.07% Indels: 74
DB: 6 Gaps: 13
US-09-208-629f-6 (1-408) x I87848 (1-1475)
Qy 18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr---- 36
Db 221 GTTCATCTGAATGTTCCATTATAAACACACAGACGCTTACTGTGGACCATTTATCTCAG 280
Qy 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
Db 281 TAATGATTCTGCTCCCGCTTTCTTTGTACAGGACGCAACACAGTAAGGAAGTCTTA 340
Qy 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
Db 341 TTGGCAGATTAGAAACCCAGCTCCATCACTACGGAAGGGTTCGGTA-GAACACAGGC 399
Qy 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThr 84
Db 400 TTTTCCATCATGATGATTC----- 417
Qy 85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
Db 418 -----TTCGCTCC----- 426
Qy 105 GlyTrpLeuThrSerSerLeuSerThrLysLeuIleProAlaIleTrpLeuLeuValPhe 124
Db 427 ---ATCCTCACCGGAAGCTGACCACGGTCTTTCTCCGGTCTCTACATTATGTGTT 483
Qy 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
Db 484 GTGATTGGTTTGCCAGTAATGTCATGGCCCTCTGGATCTCTCTTTCCGACGAAGAAG 543
Qy 145 IleCysThrThrValPheTy-----ThrAsnLeuAlaIleAlaPhePheCysVal 163
Db 544 AAACACCCCGCGGATTTACATGGCCAACTGGCCCTTGCCGACCTCTCTCTGTCATC 603
Qy 164 ThrLeuProPheLysIleAlaTrpHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
Db 604 TGGTTCCCTGGAAGATCTCTACCATCTACATGGCAACACTGGGTCTACGGGGAGGCC 663
Qy 184 LeuCysArgAlaThrThrValIlePheTyGlyAsnMetTrpCysSerIleLeuLeuLeu 203
Db 664 CTGTGCAAGGTGCTCATTTGCTTTCTATGTGTAACATGATTCCTCCATCTCTTCATG 723
Qy 204 AlaCysIleSerIleAsnArgTrpLeuAlaIleValHisProPheThrTrpArgGlyLeu 223
Db 724 ACCTGCTCAGCGGACAGAGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
Qy 224 ProLysHisThrTrpAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240

```

775 CCCAGGAAGAAG--GCAAAACATCCCGTGGCGTCTCTCTGGCAATCTGGCTCTCTGATT 831

241 PheLeuTyMetLeuProPhePheIleLeuLysGlnGluTyTrpLeuValGlnProasp 260

832 TTTTGGTCACCATCCCTTTGATGTGTCATGAAGACAGACCATCTACATCCACATGAAAC 891

261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerSerPropheGlnLeuTy 280

892 ATCACCACTGTCCAGATGG-----CTGCCTCAGAGGATGTTG 930

281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuLeuProPhe 294

931 GTGGGGGACATGTTCAATTACTTCTCTACCTGGCCATTTGGAGTCTCTCTGTTCCGGCC 990

295 ValLeuIleIleTyCysTyAlaAlaIleIleArgThrLeuAsnAla----- 310

991 CTCCTTACTGCATCTGCCTACGTGCTCATGATCAAGACGCTCCGCTCTCTCTCTGAT 1050

311 -----TyrAspHisArgTrpLeuTrpTyValLysAlaSerLeuLeuValIle 328

1051 GAACACTAGAGAACAAAGGAGGCTATCCGACTCATCATCACCCTGCTGGCCATG 1110

329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyTrp 348

1111 TACTTTCATCTGCTTCTCTAGCAACCTCTGCTGCTAGTGCATATTCTCTAATCAA 1170

349 TyrAsnAsnThrAspGlyLeuTyPheIleTyLeuIleAlaLeuCysLeuGlySerLeu 368

1171 ACCGAGAGCAGACGACGCTACCCCTCTACCTTGTGCGCCTCTGCTGCGACCTC 1230

369 AsnSerCysLeuAspProPheLeuTyTrpPheLeuMetSerLysThr---ArgAsnHisSer 387

1231 AACAGCTGCATAGACCCCTTTGCTATTTACTTTGCTCAAAAGATTTCAGGGATCAGCC 1290

BC025432 2772 bp mRNA linear ROD 07-AUG-2002

Mus musculus, coagulation factor II (thrombin) receptor-like 1,

clone MGC:29183 IMAGE:5006769, mRNA, complete cds.

BC025432

BC025432.1 GI:19343971

MG.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2772)

Strausberg, R.

Direct Submission

Submitted (06-MAR-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@nigri.nih.gov

Akher, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 39 Row: d Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

Location/Qualifiers
 1..2772

/organism="Mus musculus"
 /db_xref="taxon:10090"
 /map="FVB/N-3"
 /clone="MGC:29183 IMAGE:5006769"
 /tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month old mouse, Taken by biopsy."
 /clone_lib="NCI CGAP Mam2"
 /lab_host="DH10B"
 /note="vector: pcMV-SPORT6"
 99..1298
 /codon_start=1
 /product="coagulation factor II (thrombin) receptor-like 1"
 /protein_id="AAH25432.1"
 /db_xref="GI:19343972"
 /db_xref="LocusID:14063"
 /translation="MRSLSLAWLGGITLLAASVSCSRSTENLAFGRNNSKGRSLIGRL ETOPPTGKGVPEPGFSIDFASILITGKLTTFVWVYIIIVFVIGLPSNGMALWIF LFRKKHAPVIYMANLADLSVIFPLAIAVHLGNWVYGEALCKVLIGFFVGN MYCSILPMTCLSVQRYWVIVNMPGHPKXIANIAGVSLAIWLLIFLTVPLVYMKOTI YIPALNITTDHVLPEVLPVGMENFSLAIGVFLPAILTASAVYLMIKTRSSAM DESEKRRQAIRLITVLAWYFICFAPSLLVLLVHYFLIKTORQSHVYALYVALCL SLNSCSDIPFVYFVSKDFRDHARNALLCSVRVNRMQISLSNKFSRKSGSYSSSS TSVKTSY"

CDS

BASE COUNT 671 a 689 c 638 g 774 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,97e-40 Length: 2772
 Score: 605.50 Matches: 137
 Percent Similarity: 52.97% Conservative: 68
 Best Local Similarity: 35.40% Mismatches: 127
 Query Match: 28.07% Indels: 55
 DB: 10 Gaps: 12

US-09-208-629F-6 (1-408) x BC025432 (1-2772)

Qy 22 AlaLeuLeuPheAlaAlaAlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGly 41
 Db 111 AGCTGGCGGTGGTGGGAGGTATCACCCCTCTCTGGCGGCTCGGTCTCTGCAGCCGG 170
 Qy 42 MetGluAsn-----AspThrAsnAsnLeuAlaLysProThrLeuProIleLysThrPhe 59
 Db 171 ACCGAGAACCTTGCACCGGAGCAACACAGTAAGGAGAGAGTCTTATTTGGCAGATTG 230
 Qy 60 ArgGlyAlaProProAsnSerPheGluGluPheProPheSerAlaLeuGluGlyTrpThr 79
 Db 231 GAAACCCAGCCTCA-----ATCCTACCGGGAGCTGACCATTCTTCTTCCTCCG 251
 Qy 80 GlyAlaThrIleThrValLys-----IleLysCysProGluGluSerAlaSerHisLeu 97
 Db 252 GGGAAAGGGGTCCGGTAGAACACAGGCTTTTCCATCATGAGTCTCTCGCTCC----- 305
 Qy 98 HisValLysAsnAlaThrMetGlyThrLeuThrSerSerLeuSerThrLysLeuIlePro 117
 Db 306 -----ATCCTACCGGGAGCTGACCATTCTTCTTCCTCCG 341
 Qy 118 AlaIleThrLeuLeuValPheValIcGlyValProAlaAsnAlaValThrLeuTrpMet 137
 Db 342 GTCGCTACATTATTGTGTGTGATTTGGTGGCCAGTAATGATGCATGCCCTCTCGATC 401
 Qy 138 LeuPhePheArgThrArgSerIleCysThrThrValPheThr-----ThrAsnLeuAlaIle 156
 Db 402 TTCCTTTTCCGAGCAAGAAAGAACACCCCGCGGTGATTTACATGGCCAACTGGCGCTG 461

Qy 157 AlaAspPheLeuPheCysValThrLeuProPheLysIleAlaThrHisLeuAsnGlyAsn 176
 Db 462 GCTGACCTCTCTCTGTCTGTCTGTCCCTCCCTGGCCATTCACCTACCTACATGCAAC 521
 Qy 177 AsnTrpValPheGlyGluValLeuCysArgAlaThrThrValIlePheThrGlyAsnMet 196
 Db 522 AACTGGGTCTATGGGAGGCCCTGTGCAAGGTGCTCATTTGGCTTTTCTATGCAACATG 581
 Qy 197 TyrCysSerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHis 216
 Db 582 TATTGCTCCATCTCTTCATGACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAAC 641
 Qy 217 ProPheThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeu--- 235
 Db 642 CCCATG-----GGGCACCCAGGAAGAAG---GCAAAACATCGCGGTGGCTCTCC 689
 Qy 236 -----ValTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGlu 253
 Db 690 TTGGCAATCTGGCTCTCTGATTTTCTGGTCACCATCCCTTTGTATGTCATGAAGCAGAC 749
 Qy 254 TyrTyrLeuValGlnProAspIleThrThrCysHisAspValHisAsnThrCysGluSer 273
 Db 750 ATCTACATTCACCCCTTGAACATCACACTGCCAGATGTG----- 791
 Qy 274 SerSerProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhe 287
 Db 792 ---CTGCTGAGGAGGTATTGGTGGGACATGTTCAATTACTTCTCTCACTGGCCATT 848
 Qy 288 PheGlyPheLeuLeuProPheValLeuIleIleTyrCysTyrAlaIleIleArgThr 307
 Db 849 GGAGTCTTCTGTTCCTGGCCATCTTACTGCAATCTGCTACGTGCTCATGATCAAGACG 908
 Qy 308 LeuAsnAla-----TyrAspHisArgTrpLeuTrpTyrValLysAla 321
 Db 909 CTCGGCTCTCTCTGATGATGAACACTCAGAAAGAAAGGAGGAGGCTATCCGACTC 968
 Qy 322 SerLeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleLeuLeu 341
 Db 969 ATCATCACCGTGTGGCCATGACTTCTGCTTCTCTAGCAACCTTCTGCTGTA 1028
 Qy 342 IleHisHisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleThrLeu 361
 Db 1029 GTGCATTATTTCCTAAATCAAAACCCAGAGCAGACGCTCTACGCCCTCTACCTGCTC 1088
 Qy 362 AlaLeuCysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSer 381
 Db 1089 GCCCTCTGCTGCTGACTCTCAACAGCTGCATAGACCCCTTGTCTATTACTTTGCTCA 1148
 Qy 382 LysThr---ArgAsnHisSer 387
 Db 1149 AAAGATTTCAGGATCACGCC 1169
 RESULT 21
 MPMROACR2
 LOCUS 1477 bp DNA linear ROD 30-JAN-1995
 DEFINITION M.musculus exon 2 for proteinase activated receptor 2.
 ACCESSION Z35158
 VERSION Z35158.1 GI:559918
 KEYWORDS proteinase activated receptor; proteinase activated receptor 2.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Nysstedt.S., Emilsson.K., Wahlestedt,C. and Sundelin,J.
 TITLE Molecular cloning of a potential proteinase activated receptor
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (20), 9208-9212 (1994)
 MEDLINE 95023880
 PUBMED 7937743
 REFERENCE 2 (bases 1 to 1477)
 AUTHORS Nysstedt,S.
 TITLE Direct Submission

JOURNAL Submitted (11-JUL-1994) Sverker Nystedt, Division of Neurobiology,
The Wallenberg, Laboratory, Lund University, Soelvegatan 33A, Lund,
S-220 07, Sweden

REFERENCE 3 (bases 1 to 1477)

AUTHORS Nystedt,S., Larsson,A.K., Aberg,H. and Sundelin,J.
TITLE The mouse proteinase activated receptor 2 cDNA and gene-molecular
cloning and functional expression
J. Biol. Chem. 170 (1995) In press

FEATURES Location/Qualifiers
1..1477

source /organism="Mus musculus"
/db xref="taxon:10090"

gene /germline

1..1477

intron /gene="PAR-2"

<1..310

exon /gene="PAR-2"

/citation=[3]

311..>1477

/gene="PAR-2"

/citation=[3]

/number=1

1420..>1477

/gene="PAR-2"

BASE COUNT 341 a 413 c 321 g 402 t

ORIGIN

Alignment Scores:

Pred. No.: 3,5e-40 Length: 1477

Score: 603.50 Matches: 140

Percent Similarity: 52.25% Conservative: 69

Best Local Similarity: 35.00% Mismatches: 118

Query Match: 27.98% Indels: 74

DB: 10 Gaps: 13

US-09-208-629f-6 (1-408) x MMPROACR2 (1-1477)

QY 18 IleLysMetLysAlaLeuIlePheAlaAlaAlaGlyLeuLeuLeuLeuProThr--- 36
DB 221 GTTCATCTGAATGTTCCATTTAAACACAGCAGCCCTTACTGTGGACCATTTATCTCAG 280
QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
DB 281 TAATGATTCGTCCTCGCTTCTTTGTATACAGGACCAACAGTAAAGGAGAGTCTTA 340
QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
DB 341 TTGGCAGATTAGAAACCCAGCCTCAATCACTCGGAAAGGGTTCCGGTA-GAACCCAGGC 399
QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyThrGlyAlaThrIleThr 84
DB 400 TTTTCCATCATGATGTTCT----- 417
QY 85 ValLysIleLysCysProGluGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
DB 418 -----TCTGCGTCC----- 426
QY 105 GlyThrLeuThrSerSerLeuSerThrLysLeuIleProAlaIleThrLeuLeuValPhe 124
DB 427 ---ATCCTCACCGGAGAGCTGACCACGGTCTTTCTCCGGTCGTCTACATATTGTTT 483
QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
DB 484 GTGATTGGTTGGCCAGTAATGGCATGGCCCTCTGGATCTCTCTTTTCCGACGAGAGAG 543
QY 145 IleCysThrThrValPheThr-----ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
DB 544 AAACACCCCGCGTGATTATACATGGCCCAACTGGCCTTGGCCGACCTCTCTCTGTCATC 603
QY 164 ThrLeuProPheLysIleAlaThrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
DB 604 TGGTTCCCTCCCTGAGATCTCTACACCTATACGTGCAACAACTGGGTCTACGGGGAGCC 663

QY 184 LeuCysArgAlaThrThrValIlePheThrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
DB 664 CTGTGCAAGTCTCTATTGGCTTTTCTATGTGTACATGATGATGCTCCATCTCTCTCATG 723
QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
DB 724 ACCTGCTCAGCGTGCAGAGGTACTGGGTGATCGTGAACCCCATG-----GGACAC 774
QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTrpAlaThrVal 240
DB 775 CCCAGGAAGAAG---GCAACATCGCGTGGCGTCTCTTGGCAATCTGGCTCTCGATT 831
QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
DB 832 TTTCTGGTCACCATCCCTTTGTATGTATGTCATGAGCAGACCATCTACATCCAGCATGAAC 891
QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
DB 892 ATCACCACCTGTCCAGCATGTG-----CTGCCTCAGGAGGTATTG 930
QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
DB 931 GTGGGGGACATGTTCAATTACTTCTCTCAGTGGCCATTTGGAGTCTTCTGTTCGCGGC 990
QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
DB 991 CTCCTTACTGTCATCTGCTTACGTCTCATGATCAAGACGCTCGCTCTCTGCTATGGAT 1050
QY 311 -----TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuLeuValIle 328
DB 1051 GAACTCAGAGAAAGGAGGAGGCTATCCGACTCATCATCAGCGTGTGGCCATG 1110
QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
DB 1111 TACTTTCATCTGCTTGTCTCTAGCAACCTTCTGCTGATGATTAATTCCTTAATCAAA 1170
QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
DB 1171 ACCGAGGCGAGCCAGCTGTACGCCCTTCTACCTTGTGCGCTCTGCTGTGCGACCTC 1230
QY 369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
DB 1231 AACAGCTGCATAGACCCCTTTGTATTACTTGTCTCAAAAGATTTTCAGGATCAGCC 1290
RESULT 22
AC110876/c 199227 bp DNA linear HTG 16-FEB-2002
LOCUS Mus musculus clone RP23-176N6, WORKING DRAFT SEQUENCE, 8 unordered
DEFINITION pieces.
AC110876
VERSION AC110876.1 GI:18693418
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 199227)
AUTHORS Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 199227)
AUTHORS Martin,J., Hosseini,R., Peng,Y., Peng,Z., Rubin,E.M. and Cheng,J.F.
TITLE Direct Submission
JOURNAL Submitted (16-FEB-2002) Berkeley PGA, Lawrence Berkeley National
Laboratory, 1 Cyclotron rd., Berkeley, CA 94720, United States of
America
COMMENT Draft Sequence Produced by Berkeley PGA
Pga.lbl.gov
Center Code: PGABERK.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 2263: contig of 2263 bp in length
 2264 2363: gap of unknown length
 2364 2569: contig of 2806 bp in length
 5170 5269: gap of unknown length
 5270 13742: contig of 8473 bp in length
 13743 13842: gap of unknown length
 13843 21596: contig of 7754 bp in length
 21597 21696: gap of unknown length
 21697 37430: contig of 15734 bp in length
 37431 37530: gap of unknown length
 37531 72728: contig of 39748 bp in length
 72729 72778: gap of unknown length
 72779 129890: contig of 52512 bp in length
 129891 129990: gap of unknown length
 129991 199227: contig of 69237 bp in length.

FEATURES

Location/Qualifiers
 1. 199227
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-176N6"

BASE COUNT 55269 a 45000 c 45562 g 52676 t 720 others
 ORIGIN

Alignment Scores:

Pred. No.: 9.82e-38 Length: 199227
 Score: 603.50 Matches: 140
 Percent Similarity: 52.25% Conservative: 69
 Best Local Similarity: 35.00% Mismatches: 118
 Query Match: 27.98% Indels: 74
 DB: 2 Gaps: 13

US-09-208-629F-6 (1-408) x AC110876 (1-199227)

QY 18 IleLysMetLysAlaLeuLeuPheAlaAlaGlyLeuLeuLeuLeuProThr--- 36
 DB 174666 GTTCATCTGAAGTTCATTTAAACACAGCAGCCTTACTGTGGACCATTTATCTCAG 174607
 QY 37 -----PheCysGlnSerGlyMetGluAsnAspThrAsnAsnLeuAla----- 50
 DB 174606 TAATGATTCCTCGCTCTTCTTGTCAGCAGCAGCAACAGTAAGGAAGTCTTA 174547
 QY 51 -----LysProThrLeuProIle-----LysThrPheArgGlyAlaProPro 64
 DB 174546 TTGGCAGATTAGAAACCCAGCCTCAATCACTCGGAAGGGTTCGGTA-GAACCCAGC 174488
 QY 65 AsnSerPheGluGluPheProPheSerAlaLeuGluGlyThrThrGlyAlaThrIleThr 84
 DB 174487 TTTTCCATCATGATTC-----TCTCGCTCC----- 174470
 QY 85 ValLysIleLysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMet 104
 DB 174469 -----TCTCGCTCC----- 174461
 QY 105 GlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIleTyrLeuLeuValPhe 124
 DB 174460 ---ATCCTCACCGGGAAGCTGACACCGTCTTCTCCGGTCGTACATATTGTGTTT 174404
 QY 125 ValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSer 144
 DB 174403 GTGATTGTTTGCCAGTAATGGCATGGCCCTCTGGATCTCTTTTCCGAAGCAAG 174344
 QY 145 IleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysVal 163
 DB 174343 AAACACCCCGCGATTATACATGGCCCAACCTGGCCTTGCCGACCTCTCTGTGATC 174284
 QY 164 ThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluVal 183
 DB 174283 TGGTTCCTCCCTGAAGATCTCTACACCTACATGGCAACAACCTGGGTCTACGGGGAGCC 174224

QY 184 LeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuLeu 203
 DB 174223 CTGTGCAAGGTGCTCATTCGCTTTTCTATGTTAAACATGATTGCTCCATCTCTTCATG 174164
 QY 204 AlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeu 223
 DB 174163 ACCTGCTCCTGAGGTGACAGAGTACTGGGTGATCGTGAACCCCATG-----GGACAC 174113
 QY 224 ProLysHisThrTyrAlaLeuValThrCysGlyLeu-----ValTTPAlaThrVal 240
 DB 174112 CCAGGAAGAG---GCAACATCGCGCTTGGCGTCTCTTGGCAATCTGGCTCTCGATT 174056
 QY 241 PheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAsp 260
 DB 174055 TTTCTGCTCACCATCCCTTTGTATGTGTCATGAAGCAGACCATCTACATTCAGCATGAAC 173996
 QY 261 IleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGlnLeuTyr 280
 DB 173995 ATCACCACCTGTCCAGATGTG-----CTGCCTGAGGAGGTATTG 173957
 QY 281 -----TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPhe 294
 DB 173956 GTGGGGGACATGTTCAATTACTTCTCTCAGTGGCCATGGAGCTTCTCTGTTCCCGGCC 173897
 QY 295 ValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
 DB 173896 CTCCTTACTGTCATCTCGCTAGCTGCTCATGATCAAGACGCTCGCTCTTCTGCTATGAT 173837
 QY 311 -----TyrAspHisArgTrpLeuTyrTyrValLysAlaSerLeuLeuLeuValIle 328
 DB 173836 GAACACTCAGCAAGAAAGGAGGAGGCTATCCGACTCATCATCACCGTCTGGCCATG 173777
 QY 329 PheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyr 348
 DB 173776 TACTTCATCTGCTTCTCTCTAGCAACCTTCTCTGCTAGTGCATTATTTCCTAATCAAA 173717
 QY 349 TyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeu 368
 DB 173716 ACCCAGAGCAGAGCCACGCTACGCGCTCTACCTTGTGCGCTCTGCTGTCGACCTC 173657
 QY 369 AsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer 387
 DB 173656 AACAGCTGCATAGACCCCTTTGTCTATTACTTGTCTCAAAAGATTTCAGGGATCAGCC 173597
 RESULT 23
 XLU09632 XLU09632 1680 bp mRNA linear VRT 28-MAY-1994
 LOCUS Xenopus laevis thrombin receptor mRNA, complete cds.
 DEFINITION U09632
 ACCESSION U09632
 VERSION U09632.1 GI:495197
 KEYWORDS African clawed frog.
 SOURCE Xenopus laevis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 Xenopodinae; Xenopus.
 Xeno (bases 1 to 420)
 Gersten, R.E., Chen, J., Ishii, M., Ishii, K., Wang, L., Nanevicz, T.,
 Turk, C.W., Vu, T.K. and Coughlin, S.R.
 TITLE Specificity of the thrombin receptor for agonist peptide is defined
 by its extracellular surface
 JOURNAL Nature 368 (6472), 648-651 (1994)
 MEDLINE 94195429
 PUBMED 8145852
 REFERENCE 2 (bases 1 to 1680)
 Coughlin, S.R.
 TITLE Direct Submission
 JOURNAL Submitted (11-MAY-1994) Shaun R. Coughlin, CVRI, University of
 California, San Francisco, 505 Parnassus Avenue, Room HSW 831, San
 Francisco, CA 94143, USA
 FEATURES Location/Qualifiers
 source 1. 1680

Db	878	ATATCCATACGAGCACTATACCACCTTGTGCATCGACAACACTCAGAAGATTCTTAGG	937
QY	259	ProAspIleThrCysHisaspValHisAasnThrCysGluSerSerPropheGln	278
Db	938	CTGGATATTACTACTTGCCCATGTTTATAGAT--TTGAAGACCTCAAGACTTTTAT	994
QY	279	LeuTyTyPheIleserLeuAlaPhePheGlyPheLeuIleProPheValLeuIle	298
Db	995	ATCTACTATTCTCTTCCTTCCTGCTGCTTCTCTCGTGCCATTTCATTATCACCACT	1054
QY	299	TyrCysTyraLaalalleleargThrLeuAasAlaTyraSp-----	312
Db	1055	ATCTGCTACATAGGATCATCAGAAGCTTAAGTTCTCCAGCATCGAGAACAGTTCGAAG	1114
QY	313	--HisArgTrpLeuTrpTyraVallLysAlaSerLeuLeuLeuValIlePheThrIle	331
Db	1115	AGACAAGAGCCCTGTIC-----TTGGCGGTGGTGTGCTGTCTGTGTTTTATCATC	1165
QY	332	CysPheAlaProSerAsnIleLeuIleHISHisAlaAasnTyrtTyraAsnAsn	351
Db	1166	TGTTTTGGGCCAACCAATGCTCTTTCTGACCCAT-----TACTTGCAAGAAGCC	1216
QY	352	ThrAspGlyLeuTytyrPheIleTyLeuIleAlaLeuCysLeuGlySerLeuAasnSerCys	371
Db	1217	AATGAATTTCATATATTTGCTTACATTCAGTGCCTGTGTGGCAGCGTCAGCTGTTC	1276
QY	372	LeuAspProPheLeuTytyrPheLeuMetSer	381
Db	1277	CTCGATCCCCCTTATTACTATGCGTCA	1306
RESULT 24			
RNU61373			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			


```

/clone.lib="NIH MGC 66"
/lab_host="DH108"
/notes="Vector: pCMV-SPORT6"
CDS
93..1286
/codon_start=1
/product="Unknown (protein for MGC:9549)"
/protein_id="AAH12453.1"
/db_xref="GI:15214650"
/translation="MRSPRAWLLGAAIILAAASLSCSTIQCTNRSSKRSGLIGKVDG
RTHKGPVIVETVSVDFSAVLTKUTTVPLPIVYIIVFVGLPNSGMALWVPLF
RTKKGPAIVANALADLLSVIFPLKIAIYHGNWNIYGEALCNVLIGFFYGNMY
CSILFTCLSVORYWYIVNPMGHSRKRKANIAGISLAIWLILLVITPIYVVKQTFI
PALNITTDHVLPEQLQVDMFNFLSLAIGVFLPFLTAGAYLVIMIRMLRMSAMDE
NSPKKRAIKLIVTVLAVLILCFTPSNLLLVVHFLIKSQGSHVYALYIVALCLST
LNSCIDPFVYFVSHDFRDAKNALLCRSVIRVQKQVSLTSKHSKRKSSYSSSSTT
VKTSY"
BASE COUNT      757 a   614 c   604 g   838 t
ORIGIN

Alignment Scores:
Pred. No.:      5,76e-39      Length:      2813
Score:          592.50        Matches:     142
Percent Similarity: 50.85%    Conservative: 67
Best Local Similarity: 34.55%  Mismatches:  145
Query Match:     27.47%      Indels:      57
DB:              9           Gaps:       14

US-09-208-629F-6 (1-408) x BC012453 (1-2813)

QY      8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db      84 TCCAGGAGGATCGGAGCCCGCGGCGTGTCTGGGGCGCCGATCTCTGCTAGCA 143
QY      28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db      144 GCCTCTCTC-----TCCTGC---AGTGGCACCATCCCAAGAACCAAT 182
QY      48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProAsnSerPhe 67
Db      183 AGATCTCTAAAGAGAACGCTATTGGTAAGTTGATGGACA-----227
QY      68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValLysIle 87
Db      228 -----TCCACGTCACCTGGA---AAAGGAGTTACAGTTGAAGAACAGCTTT 269
QY      88 LysCysProGluSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db      270 TCTGTGGATGAGTTTCTGCATCT-----GTCCCTC 299
QY      108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValValGly 127
Db      300 ACTGGAAACGACCACTGCTCTCTCCCAATTGCTACCAATTGTTGTGGTGGGT 359
QY      128 ValProAlaAsnAlaValThrLeuTrpMetLeuPhePheArgThrArgSerIleCysThr 147
Db      360 TTGCCAAGTAACGCGCTGCGCTGTGGTCTTTCTTTTCCGAACCTAAGAAAGACACCC 419
QY      148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db      420 GCTGTGATTACATGGCCCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGGTTCCTCC 479
QY      167 PheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db      480 TTGNAGATTGCTATCATACATACATGCGCAACACTGGATTTATGGGAGCTCTTTGTAAT 539
QY      187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db      540 GTGCTATTGGCTTTTCTATGGCAACATGACTGTTCCATCTCTCTCATGACCTGCTC 599
QY      207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db      600 AGTGTGAGAGGATTGTTGGTTCATCGTAGAACCCCATGGGGCAC---TCCAGGAAGAAGCA 656

```

```

227 ThrTyrAlaLeuValThrCysGlyLeuValTrpAlaThrValPheLeuTyrMetLeuPro 246
Db      657 AACATTGGCATTGGCATCTCCCTGCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 716
QY      247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db      717 TTGTATGTGCTGAAGCAGACCATCTTCATTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
QY      267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr-----280
Db      777 GTT-----TTGCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 815
QY      281 TyrPheIleSerLeuAlaPheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db      816 TACTTCTCTCTCTGCGCATTTGGGTCTTTCTGTTCCCGACCTCTCTCAGACCTCTGCC 875
QY      301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db      876 TATGTCTGATGATCAGATGCTGCGATCTTCTGCGCATGGATGAAAACTCAGAGAAGAAA 935
QY      315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleLeuValIlePheThrIleCysPheAla 334
Db      936 AGGAAGAGGCGCATCAAACTATTGTACTGCTGCGCCATGACTGCTGCTGCTGCTGCTGCT 995
QY      335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrAsnAsnThrAspGly 354
Db      996 CCTAGTAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
QY      355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db      1047 CAGAGCCATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1106
QY      372 LeuAspProPhePheLeuLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db      1107 ATCGACCCCTTGTCTATTACTTTGTTTCACATGATTTTCAGGATGATGCAAGAAGAGCT 1166
QY      390 TyrLeuThrLys*****AsnAspLeuArgGlu 400
Db      1167 CTCCTTTGCGAAGTGTCCGCACTGTAAAGCAG 1199

AC128225      88955 bp      DNA      linear      HTG 31-JUL-2002
Rattus norvegicus clone CH230-498A19, *** SEQUENCING IN PROGRESS
*** 44 unordered pieces.
AC128225      AC128225.1 GI:21908835
HTG; HTGS_PHASE1.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 88955)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-ouman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homes,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

```

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,X., Liu,W., Loulesed,H.,
Lozad,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Mahehwar,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sleson,I.,
Sodergren,E., Sonaikae,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,O.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 88955)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 88955)
Worley,K.C.
Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZLT
Center clone name: CH230-498A19
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 35763 bases at least Q40
Consensus quality: 40497 bases at least Q30
Consensus quality: 44674 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1023: contig of 1023 bp in length
* 1024: gap of unknown length
* 1124: contig of 1035 bp in length
* 2158: gap of unknown length
* 2259: contig of 1138 bp in length
* 3396: gap of unknown length
* 3496: contig of 1263 bp in length
* 3497: gap of unknown length
* 4859: gap of unknown length
* 4860: contig of 1462 bp in length
* 6321: gap of unknown length
* 6422: gap of unknown length
* 7471: contig of 1050 bp in length
* 7472: gap of unknown length
*
* 7572: 8781: contig of 1210 bp in length
* 8882: 8881: gap of unknown length
* 10239: 10238: contig of 1357 bp in length
* 10339: 10338: gap of unknown length
* 11514: 11514: contig of 1176 bp in length
* 11614: 11614: gap of unknown length
* 13207: 13207: contig of 1593 bp in length
* 13308: 13307: gap of unknown length
* 14715: 14714: contig of 1407 bp in length
* 14815: 14814: gap of unknown length
* 16170: 16170: contig of 1356 bp in length
* 16270: 16270: gap of unknown length
* 17527: 17526: contig of 1256 bp in length
* 17627: 17626: gap of unknown length
* 19027: 19027: contig of 1401 bp in length
* 19128: 19127: gap of unknown length
* 20259: 20258: contig of 1131 bp in length
* 20359: 20358: gap of unknown length
* 21563: 21562: contig of 1204 bp in length
* 21663: 21662: gap of unknown length
* 23426: 23425: contig of 1763 bp in length
* 23526: 23525: gap of unknown length
* 24838: 24837: contig of 1312 bp in length
* 24938: 24937: gap of unknown length
* 25947: 25946: contig of 1009 bp in length
* 26047: 26046: gap of unknown length
* 27584: 27583: contig of 1537 bp in length
* 27684: 27683: gap of unknown length
* 29344: 29343: contig of 1660 bp in length
* 29444: 29443: gap of unknown length
* 30865: 30864: contig of 1421 bp in length
* 30965: 30964: gap of unknown length
* 32342: 32341: contig of 1377 bp in length
* 32442: 32441: gap of unknown length
* 34493: 34492: contig of 2051 bp in length
* 34593: 34592: gap of unknown length
* 37028: 37027: contig of 2435 bp in length
* 37128: 37127: gap of unknown length
* 40024: 40023: contig of 2896 bp in length
* 40124: 40123: gap of unknown length
* 42441: 42440: contig of 2317 bp in length
* 42541: 42540: gap of unknown length
* 45194: 45193: contig of 2653 bp in length
* 45294: 45293: gap of unknown length
* 47983: 47982: contig of 2689 bp in length
* 48083: 48082: gap of unknown length
* 49559: 49558: contig of 1376 bp in length
* 51559: 51558: gap of unknown length
* 51659: 51658: gap of unknown length
* 53032: 53031: contig of 1373 bp in length
* 53132: 53131: gap of unknown length
* 55704: 55703: contig of 2572 bp in length
* 55804: 55803: gap of unknown length
* 57503: 57502: contig of 1699 bp in length
* 57603: 57602: gap of unknown length
* 59339: 59338: contig of 2336 bp in length
* 60039: 60038: gap of unknown length
* 62415: 62414: contig of 2276 bp in length
* 62515: 62514: gap of unknown length
* 65537: 65536: contig of 3122 bp in length
* 68901: 68900: contig of 3265 bp in length
* 69002: 69001: gap of unknown length
* 71415: 71414: contig of 2413 bp in length
* 74439: 74438: gap of unknown length
* 74540: 74539: contig of 2925 bp in length
* 78489: 78488: gap of unknown length
* 78590: 78589: contig of 3950 bp in length
* 82299: 82298: gap of unknown length
* 85743: 85742: contig of 3610 bp in length
* 88200: 88199: contig of 3444 bp in length

```

TITLE
REFERENCE
JOURNAL
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

Waldron, L., Walker, B.,

Waldron, L., Walker, B.,


```
Db 95119 CTCCTCTCTGTCATCTGGTTCCCTCCAGAGATCTCTACACCTCCATGCAACGACTGG 35060
Qy 179 valphegylcluvallleuCyArgAlaThrValIlePheThrValIlePheThrGlyAsnMetTyrCys 198
Db 95059 ACCTATGGGATCGCTCTGCAAGGTCTCATTTGGCTTTTCTACGGCAATATGTACTGC 95000
Qy 199 SerileLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db 94999 TCCATCTCTTCATGACCTGCTCAGGTGCAGGTACTGGGTGATCTGTGAACCCATG 94940
Qy 219 ThrTyrArgGlyLeuProIlyHisThrTyrAlaLeuValThrCysGlyLeu----- 235
Db 94939 GGACAC-----TCCAGGAAGAGGGCAACATCGCTGCTGGCTCTCCCTGGCC 94892
Qy 236 valTrpAlaThrValPheLeuTyrMetLeuProPhePheIleLeuIysGlnGluTyrTyr 255
Db 94891 ATCTGGCTCTGATTTTCTGGTCCACATCCCTCTGTACGTATGAGGAGACCATCTAC 94832
Qy 256 LeuValGlnProAspIleThrCysHisAspValHisAsnThrCysGluSerSerSer 275
Db 94831 ATTCAGACCTTGACATCACCATCTGTACGACGTG-----CTG 94793
Qy 276 ProPheGlnLeuTyr-----TyrPheIleSerLeuAlaPhePheGly 289
Db 94792 CCGGAGGAGTCTCTGGTGGGGACATGTTTCAGTTACTCTCTCTCCCTGGCCATTGGAGTC 94733
Qy 290 PheLeuLeuProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsn 309
Db 94732 TTTCTGTCTCCAGCCCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 94673
Qy 310 Ala-----TyrAspHisArgTrpLeuTrpTyrValIysAlaSerLeu 323
Db 94672 TCTCCGCCATGACGAGCACTCGGAGAGAAAGCGGGGCTATCCGCTCATCATC 94613
Qy 324 LeuLeuLeuValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleHis 343
Db 94612 ACGGTGTCGTCATCTACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 94553
Qy 344 HisAlaAsnTyrTyrTyrAsnAsnThrAspGlyLeuTyrPheIleTyrLeuIleAlaLeu 363
Db 94552 TATTTCTCATCAAAAGCCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 94493
Qy 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db 94492 TGCTGTCCACCTCAACAGCTGATAGACCCCTTTGTCTACTACTTTGTTTCGAAA--- 94436
Qy 384 ArgAsnHisSerThrAlaTyrLeuThrLys*****AsnAspLeuArgGluGlnGlyGln 403
Db 94435 -----GATTTTCAGGAGCAGGCCAGA 94415
```

```
RESULT 28
BC018130 2876 bp mRNA linear PRI 06-DEC-2001
LOCUS Homo sapiens, coagulation factor II (thrombin) receptor-like 1,
DEFINITION clone MGC:9298 IMAGE:3895653, mRNA, complete cds.
ACCESSION BC018130
VERSION BC018130.1 GI:17390291
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2876)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
```

Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 14 Row: h Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 8051581.

FEATURES

Location/Qualifiers
1..2876
/organism="Homo sapiens"
/db_xref="LocusID:2150"
/db_xref="taxon:9606"
/clone="MGC:9298 IMAGE:3895653"
/tissue_type="Pancreas, epithelioid carcinoma"
/clone_lib="NIH_MGC_70"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
155..1348
/codon_start=1
/product="coagulation factor II (thrombin) receptor-like
1"
/protein_id="AAH18130.1"
/translation="MRSPSAAMLGAAILAASLCSGTIOGTNRSSKGRSLIGKVDD
TSHVTKGVTVETVFSVDFASVLTGKLTTVFLVTVTVFVVGLPNSGMALVFLF
RTKKHPAVIYMANLADLLSVIFPLKIAVHIGNNWIYGEALCNVLIGFYGNMY
CSILFMTCLSVORYVIVNPMGHSRKKANIGISLAWLLILLVTIPIYVVKQTIFI
PALNITTDHVLPEQLVGMDFNFIPLSIAIGVFLFPAPFATASAYVLMIRLSSANDE
NSEKRRKALIKIVSLAMLYITCTPSNLLVYVFLIKSQGSHVYALYIIVALCLST
LNSCIDPFVYVSHDFRDKHAKNLLCSRVTKQVMSLTSKHSRKSSTSSSTT
VNTSY"

CDS

BASE COUNT 771 a 632 c 629 g 844 t
ORIGIN
Alignment Scores:
Pred. No.: 7,13e-39 Length: 2876
Score: 591.50 Matches: 142
Percent Similarity: 50.85% Conservatives: 67
Best Local Similarity: 34.55% Mismatches: 145
Query Match: 27.42% Indels: 57
DB: 9 Gaps: 14

US-09-208-629F-6 (1-408) x BC018130 (1-2876)

```
Qy 8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaIle 27
Db 146 TCCAGGAGGATCGGAGCCGCCGCGTCTGCTGGGGCGGCCCATCTCTGTAGCA 205
Qy 28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db 206 GCCTCTCTC-----TCTTGC---AGTGGCACCATTCCAGGAACCAAT 244
Qy 48 AsnLeuAlaLysProThrLeuProIleLysThrPheArgGlyAlaProAsnSerPhe 67
Db 245 AGATCTCTTAAGGAAGAGCCTTATTTGTTAGGTTGATGGCACA----- 289
Qy 68 GluGluPheProPheSerAlaLeuGluGlyTrpThrGlyAlaThrIleThrValIle 87
Db 290 -----TCCACGCTCACTGGA---AAAGGAGTTACAGTTGAACAGCTCTTT 331
Qy 88 LysCysProGluCysSerAlaSerHisLeuHisValLysAsnAlaThrMetGlyTyrLeu 107
Db 332 TCTGTGGATGAGTTTCTGCATCT-----GTCCCTC 361
Qy 108 ThrSerSerLeuSerThrLysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
```

```

Db      362 ACTGGAAACTGACCACTGCTCTCTCCATATGCTACACAATGGTTGGTGGGGT 421
Qy      128 ValProAlaAsnAlaValThrLeuTyrMetLeuPhePheArgThrArgSerIleCysThr 147
Db      422 TTGCCAAGTAACGGCATGCCCTGTGGTCTTCTTTCCGAACCTAAGAGAAGACCCT 481
Qy      148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db      482 GCTGTGATTACATGGCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGTTCCCC 541
Qy      167 PheIysIleAlaTyrHisLeuAsnGlyAsnAsnTyrValPheGlyGluValLeuCysArg 186
Db      542 TTGAAGATTCCCTATCATACATATGCAACCACTGGATTATGGGAAGCTCTTTGTAAT 601
Qy      187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db      602 GTGCTATTGGCTTTTCTATGGCAACATGATGCTACTGTTCATCTCTCATGACCTGCTC 661
Qy      207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProIysHis 226
Db      662 AGTGTGCAGAGGTATTGGGTATGCTGCAACCCCATGGGCAC---TCCAGGAAGAAGCA 718
Qy      227 ThrTyrAlaLeuValThrCysGlyLeuValTyrAlaThrValPheLeuTyrMetLeuPro 246
Db      719 AACATTGCCAATGGCACTCCCTGGCAATATGGCTGCTGATTCTGCTGGTCAACATCCCT 778
Qy      247 PhePheIleLeuIysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db      779 TTGTATGCTGTGAAGCAGACCATCTTCATCTCCGCCCTGAACTCAGCATCAGCATGTCAT 838
Qy      267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr----- 280
Db      839 GTT-----TTGCTTGAGCAGCAGCTCTTGGTGGGAGACATGTTCAAT 877
Qy      281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db      878 TACTTCTCTCTCTGGCCATTTGGGCTTCTTCTGTTCCCAAGCTTCTCCTCAGACCTCTGCC 937
Qy      301 TyrAlaAlaIleIleAlaGThrLeuAsnAla-----TyrAspHisArg 314
Db      938 TATGTGCTGATGATCAGAAATGCTGCGATCTTCTGCCATGGATGAAACTCAGAGAAGAAA 997
Qy      315 TrpLeuTyrTyrValIysAlaSerLeuLeuLeuValIlePheThrIleCysPheAla 334
Db      998 AGGAAGAGGGCCATCAAACTCATTTGCTCTGCTGCTGCCATGATACCTGATCTGCTCACT 1057
Qy      335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db      1058 CCTAGTAACCTTCTGCTTGTGGTGCAT-----TATTTTCTGATTAAGAGCCAGGCG 1108
Qy      355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db      1109 CAGAGCCATGCTTATAGCCCTGTACATGTAGCCCTCTGCTCTCTACCCCTTAACAGCTGC 1168
Qy      372 LeuAspProPheLeuTyrPheLeuMetSerIysThr---ArgAsnHisSer---ThrAla 389
Db      1169 ATCGACCCCTTTGTCTATTACTTTGTTTTCATGATTTTCAGGATCATGCAAGAAGCGCT 1228
Qy      390 TyrLeuThrIys*****AsnAspLeuArgGlu 400
Db      1229 CTCCTTTTGGCGAAGTGTCCGCACTGTAAAGCAG 1261

RESULT 29
HSU34038
LOCUS      1451 bp mRNA linear PRI 21-FEB-1997
DEFINITION Human proteinase-activated receptor-2 mRNA, complete cds.
ACCESSION U34038
VERSION    U34038.1 GI:1041728
KEYWORDS
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 1451)
Bohm, S.K., Kong, W., Bromme, D., Smeekens, S.P., Anderson, D.C.,
Connolly, A., Kahn, M., Nelken, N.A., Coughlin, S.R., Payan, D.G. and
Bunnett, N.W.
Molecular cloning, expression and potential functions of the human
proteinase-activated receptor-2
Biochem. J. 314 (Pt 3), 1009-1016 (1996)
96177879
PUBMED     8615752
REFERENCE  2 (bases 1 to 1451)
AUTHORS    Bohm, S.
TITLE       Direct Submission
JOURNAL     Submitted (15-AUG-1995) Stephan Bohm, Department of Surgery, School
of Medicine, University of California at San Francisco, 521
Parnassus Avenue, San Francisco, CA 94143-0660, USA
FEATURES
source      1. .1451
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /tissue_type="kidney"
            1. .147
            /note="G-protein-coupled receptor"
            148. 1341
            /product="proteinase-activated receptor-2"
            /protein_id="AAB47871.1"
            /db_xref="GI:1041729"
            /translation="MRSPSAWMLGAAIILAAASLCSGTIOGTNRSSKGRSLIGKVDG
            TSHVTKGVTVEVSVDFSAVLTKLTTPVPIVYIVFVVGPGPSMGMAWMLVLF
            RTKKHPAVIYMANIALADLLSVFPLKIAIHANHWIYGEALCNVLIGFFIGNMY
            CSILFWTCSVQRYVYVNPMSHRKKAIAIGISAIWLILLVLTPIYVVKOTIFI
            PALNTTCHDVLPEQLLVGMENYFLSAIGVLPFAFLTASAYVIMRLRSSAMDE
            NSEKKRAIKLIVTLAMLYICTPSNLLLVVHPLIKSQOSHVVYALYIVALCLST
            LNSCIDPPVYVYFVSHDFRDHAKNALLCRSVRTKQMVSLTSKKHSKSSSYSSSTT
            VKTSY"
3'UTR      1342. .1451
BASE COUNT 310 a 389 c 346 g 406 t.
ORIGIN
Alignment Scores:
Pred. No.: 8'29e-39 Length: 1451
Score: 586.50 Matches: 141
Percent Similarity: 50.61% Conservative: 67
Best Local Similarity: 34.31% Mismatches: 146
Query Match: 27.19% Indels: 57
DB: Gaps: 14
US-09-208-629F-6 (1-408) x HSU34038 (1-1451)
Qy      8 Ser***ArgLeuArgAspGlyThrGlnValIleLysMetLysAlaLeuIlePheAlaAla 27
Db      139 TCCAGAGGATGCGGAGCCCGCGGCGTGGTCTGCTGGGGCCCGCATCTCTGTAGCA 198
Qy      28 AlaGlyLeuLeuLeuLeuLeuProThrPheCysGlnSerGlyMetGluAsnAspThrAsn 47
Db      199 GCCTCTCTC-----TCTGTC---AGTGGCACCATTCCAAGAACCAAT 237
Qy      48 AsnLeuAlaIysProThrLeuProIleLysThrPheArgGlyAlaProProAsnSerPhe 67
Db      238 AGATCTCTTAAAGGAAGACCTTATTTGGTAAGGTTGATGGGCACA----- 282
Qy      68 GluGluPheProPheSerAlaLeuGluGlyTyrThrGlyAlaThrIleThrValIle 87
Db      283 -----TCCACGCTCACTGGA---AAAGGAGTTACAGTTGAACAGCTCTTT 324
Qy      88 LysCysProGluGluSerAlaSerHisLeuHisValIysAsnAlaThrMetGlyTyrLeu 107
Db      325 TCTGTGGATGATGTTTCTGTCATCT-----GTCTC 354
Qy      108 ThrSerSerLeuSerThrIysLeuIleProAlaIleTyrLeuLeuValPheValGly 127
Db      355 ACTGGAAACTGACCACTGCTCTCTTCCAAATGTTCTACACAATTTGTTGGTGGGT 414

```

```
QY 128 ValProAlaAsnAlaValThrLeuTrpMetLeuPheArgThrArgSerIleCysThr 147
Db 415 TTGCCAAGTAACGGCATGGCCCTGTGGTCTTTCTTCCGAAGTAAGAGAGACCCCT 474
QY 148 ThrValPheTyr---ThrAsnLeuAlaIleAlaAspPheLeuPheCysValThrLeuPro 166
Db 475 GCTGTGATTTACATGGCAATCTGGCTTGGCTGACCTCTCTCTGTCATCTGTTTCCCC 534
QY 167 PheIleAlaIleTyrHisLeuAsnGlyAsnAsnTrpValPheGlyGluValLeuCysArg 186
Db 535 TTGAAGATTCCCTATCATCATACATATGCAACAACTGGATTTATGGGAAGCTCTTTGTAAT 594
QY 187 AlaThrThrValIlePheTyrGlyAsnMetTyrCysSerIleLeuLeuAlaCysIle 206
Db 595 GTGCTATTGGCTTTTCTATGGCAACATGATGCTGATTTCTTCATGACCTGCTC 654
QY 207 SerIleAsnArgTyrLeuAlaIleValHisProPheThrTyrArgGlyLeuProLysHis 226
Db 655 AGTGTGCAGAGGATTGGGTTCATCGTGAACCCCATGGGCAC---TCCAGGAAGAAGGCA 711
QY 227 ThrTyrAlaLeuValThrCysGlyLeuValThrAlaThrPheLeuTyrMetLeuPro 246
Db 712 AACATTGCCATTGGCAATCTCCCTGGCAATATGCTGCTGATTTCTGTGGTCACCATCCCT 771
QY 247 PhePheIleLeuLysGlnGluTyrTyrLeuValGlnProAspIleThrThrCysHisAsp 266
Db 772 TTGTATGTCGTGAAGCAGACATCTTCATTCCTGCTGACATCAGCATCTGTCATGAT 831
QY 267 ValHisAsnThrCysGluSerSerProPheGlnLeuTyr-----TyrAspHisArg 314
Db 832 GTT-----TTGCCCTGAGCAGCTCTTGGTGGGAGACATGTTTCAAT 870
QY 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValLeuIleIleTyrCys 300
Db 871 TACTTCTCTCTCTGGCCATTTGGGCTCTTCTGTTCCTCCAGCTTCTCCTCAGCCTCTGCC 930
QY 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 931 TATGTGCTGATGATCAGATGCTGCGATCTTCCCATGGATGAAACTCAGAGAAGAA 990
QY 315 TrpLeuTrpTyrValLysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
Db 991 AGGAAGAGGGCCCAACTCATTTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050
QY 335 ProSerAsnIleIleLeuIleIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 1051 CCTAGTAACCTTCTGCTTGGTGGCAT-----TATTTCTGATTAAAGCCAGGGC 1101
QY 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1102 CAGAGCCATGCTATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
QY 372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db 1162 ATGACCCCTTTGCTATTACTTTGTTTTCATGATTTTACGATTTACGATCATGCAAGAAGCT 1221
QY 390 TyrLeuThrLys*****AsnAspLeuArgGlu 400
Db 1222 CTCCTTTGCCGAAGTGTCCGACATGTAAGACAG 1254
RESULT 30
HSU36753
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1124)
1124 bp DNA linear PRI 28-FEB-1996
Human protease-activated receptor 2 gene, exon 2 and partial cds.
U36753.1 GI:1208539
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1124)
Kahn, M.L. and Coughlin, S.R.
Direct Submission
Submitted (19-SEP-1995) Mark L. Kahn, CVRI/Medicine, UCSF, 3rd and
Parnassus, San Francisco, CA 94143-0524, USA
FEATURES
Location/Qualifiers
1..1124
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5q13"
/note="Human genomic DNA in P1 bacteriophage library"
<3..1112
/codon_start=1
/product="protease-activated receptor 2"
/protein_id="AA90957.1"
/db_xref="GI:1208540"
/translation="TNRSGKSLGKVDGTSHTVGVTVTFVSDFSVASLVTKG
LTFVPLPIYTVFVGVFSGMGLWFLRTKKHPAVIYMANLADLLSVIFPL
KIAYHGNMNIYVGVFSGMGLWFLRTKKHPAVIYMANLADLLSVIFPL
NIAIGSLAIWLLIILLVTPVVKQITFI PALNITDCHDLVPEOLLVDMFNYPULS
AIGVLEPFLATASAVHYMIRLSAMDENSEKKRAIKLIVTLVLAALICFTPSN
LLLVHYFLIKSQGSHVYALIVLALICLSLNSCIDPFVYFVSHDFRPHAKNALLCR
SVRTVKQWQSLTSKKHRSKSSSYSSSTTVTKTSY"
3..1110
/product="protease-activated receptor 2"
/number=2
BASE COUNT 255 a 292 c 238 g 339 t
ORIGIN
Alignment Scores:
Pred. No.: 7,45e-39 Length: 1124
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: Gaps: 10
US-09-208-629f-6 (1-408) x HSU36753 (1-1124)
QY 80 GlyAlaThrIleThrValLysIleLysCysProGluSerAlaSerHisLeuHisVal 99
Db 72 GGAGTTACAGTTGAAACAGCTCTTTCTGTGGATGAGTTTTCGCATCT----- 119
QY 100 LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle 119
Db 120 -----GTCCTCACTGGAAACTGACCCTGCTCTCTCTTCCCAATTGTC 161
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 162 TACACAATTGTTGTTGGTGGTTTGCCCAAGTAACGGCATGGCCCTGTTGGTCTTCTT 221
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 222 TTCCGAACCTAAGAAAGAACGACCCCTGCTGATTTACATGCGCAATCTGCGTTGGCGGAC 281
QY 159 PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTyr 178
Db 282 CTCCTCTCTGTCATCTGGTTCCCTTGAAGATTGCTATCATACATACATGCAACAACCTGG 341
QY 179 ValPheGlyGluValLeuCysArgAlaThrValIlePheTyrGlyAsnMetTyrCys 198
Db 342 ATTATGGGGAAGCTCTTTGTAATGCTTATTGGCTTTTCTATGGTAACATGTAAGTCTGT 401
QY 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProphe 218
Db 402 TCCATTCTCTTACCTGACCTGCTGATGTCGAGAGGATTTGGGTTCATCTGCAACCCCATG 461
QY 219 ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTyrAla 238
Db 462 GGGCAC---TCCAGGAGAGGCAACATTTGCGATGCGATCTCCCTGGCAATATGGCTG 518
QY 239 ThrValPheLeuTyrMetLeuProPheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db 239 ThrValPheLeuTyrMetLeuProPheIleLeuLysGlnGluTyrTyrLeuValGln 258
```

519	CTGATTCTTCGGTCACCATCCCTTTGATGTGCGAAGCAGACCATTCTTCATCTCTGCC	578
259	ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln	278
579	CTGAACATCAGCACCTGCCATGATTT	617
279	LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle	292
618	CTCTGGTGGGAGACATGTTCAATTACTTCTCTCTCTGGCCATTGGGGTCTTTCGTTC	677
293	ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAenAla----	310
678	CCAGCCTTCCTCACAGCCTCTGCCTATGTGCTGATGATCAGAATGCTCGCATCTTCGCC	737
311	-----TyrAspHisArgTrpLeuTrpTyrValIysAlaSerLeuLeuLeuLeu	326
738	ATGGATGAAACTCAGAGAAGAGGAGGCCATCAAACTCATGTGCTATGTCCTCG	797
327	ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn	346
798	GCCATGTACCTGATCTGCTTCACCTCTAGTAACCTTCTGCTTGGTGGCAT-----	848
347	TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu	363
849	TATTTTCTGATTAAAGCCAGGCCAGAGCCATGCTATGCCCTGTGATCATTTGAGGCCCTC	908
364	CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr	383
909	TGCCCTCTACCTTAAACAGCTGCATCAGCCCTTTGCTATTACTTTGTTTCACATGAT	968
384	----ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AenAspLeuArgGlu	400
969	TTCAGGGATCATGCAAGAAACGCTCTCCTTTGGCGAAGTGTCCGCACTGTAAAGACG	1025

RESULT 31

AKU12638	LOCUS	AR012638	1255 bp.	DNA	linear	PAT 04-DEC-1998
	DEFINITION	Sequence 3	from patent	US 5763575.		
	ACCESSION	AR012638				
	VERSION	AR012638.1	GI:3970628			

BASE COUNT	294 a	320 c	260 g	381 t
ORIGIN				

Alignment Scores:

Argument Scores:		
Pred. No.:	8.46e-39	Length: 1255
Score:	585.50	Matches: 124
Percent Similarity:	54.87%	Conservative: 62
Best Local Similarity:	36.58%	Mismatches: 114
Query Match:	27.14%	Indels: 39
DB:	6	Gaps: 10

US-09-208-629F-6 (1-408) x AR012638 (1-1255)

Qy	80	GlyAlaThrIleThrValLysIleLysCysProGluSerAlaSerHisLeuHisVal	99
Db	212	GGAGTTACAGTTGAAACAGCTCTTTCTGTGATGAGTGTCTTGCATCT	259
Qy	100	LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuIleProAlaIle	119
Db	260	-----GTCCTCACTGGAAAACTGACCCTGTCTCTCTTCCATTTGTC	301
Qy	120	TyrLeuLeuValpheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe	139

Qy	239	ThrValPheLeuTyrMetLeuProPhePheLeuLysGlnGluTyrTyrLeuValGln	258
Db	659	CTGATTCGTGGTGCACCATCCCTTTGTATGTCGTGAAGCACCATTCTTCATCTCCTGCC	718
Qy	259	ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerSerProPheGln	278
Db	719	CTGAACATCAGCAGCTGTCATGATGTT-----TTGCCTAGACGAG	757
Qy	279	LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle	292
Db	758	CTCTGTGGTGGAGACATGTTCAATTACTTCTCTCTCGCCATTGGGGCTTTCGTCTTC	817
Qy	293	ProPheValleuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----	310
Db	818	CCAGCCTTCCTACAGAGCTCTGCCTATGTGCTGATGATCAGAAATGCTCGCATCTTCTGCC	877
Qy	311	-----TyrAspHisArgTyrLeuTyrTyrVallysAlaSerLeuLeuLeuLeu	326
Db	878	ATGATGAAAACTCAGAGAGAAAGAAAGAGGGCCATCAAACTCATGTGCATGTCCTG	937
Qy	327	ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn	346
Db	938	GCCATGACCTGATCTGCTTCACTCTAGTAACCTTCTGCTGTGGTGTCAT-----	988
Qy	347	TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuLeuAlaLeu	363
Db	989	TATTTTCTGATTAAAGACCGAGGGCCAGAGCCATGCTATGCGCTGTACATTGTAGCCCTC	1048
Qy	364	CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr	383
Db	1049	TGCTCTCTACCCCTTAACAGCTGCATCGACCCCTTTGTCTATTACTTTGTTTACATGAT	1108
Qy	384	--ArgAsnHisSer--ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu	400
Db	1109	TTCAGGATCATGCAAGACGCTCTCTTGGCGAAGTGTCCGCACCTGTAAGACGAG	1165
RESULT 34			
LOCUS	I87849	1255 bp	DNA linear PAT 10-AUG-1998
DEFINITION	Sequence 3 from patent US 5716789.		
ACCESSION	I87849		
VERSION	I87849.1	GI:3407789	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 1255)		
AUTHORS	Sundelin,J. and Scarborough,R.M.		
TITLE	Method to determine ligands, agonist and antagonist of C140 receptor		
JOURNAL	Patent: US 5716789-A 3 10-FEB-1998;		
FEATURES	Location/Qualifiers		
source	1..1255		
BASE COUNT	294 a 320 c 260 g	381 t	
ORIGIN	/organism="unknown"		
Alignment Scores:			
Pred. No.:	8.46e-39	Length:	1255
Score:	585.50	Matches:	124
Percent Similarity:	54.87%	Conservative:	62
Best Local Similarity:	36.58%	Mismatches:	114
Query Match:	27.14%	Indels:	39
DB:	6	Gaps:	10
US-09-208-629F-6 (1-408) x I87849 (1-1255)			
Qy	80	GlyAlaThrIleThrVallysIleLysCysProGluGluSerAlaSerHisLeuHisVal	99
Db	212	GGAGTTACAGTTGAACAGCTTTTCTGTGATCAGTTTCTGCATCT-----	259
Qy	100	LysAsnAlaThrMetGlyTyrLeuThrSerSerLeuSerThrLysLeuLeuProAlaIle	119

Db	260	-----GTCCTCACTGGAAACTGACCACATGTCTTCCCTCCAATTGC 301
Qy	120	TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db	302	TACACAATTGTTTGTTGGTTGGTTGCCAGGTAAAGCGCATGGCCTGTGGGTCTTCTT 361
Qy	140	PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db	362	TTCCGAACCTAAGAAGAAGCACCCCTGCTGTGATTTACATGGCCAATCTGGCGCTTGGCTGAC 421
Qy	159	PheLeuPheCysValThrLeuProPheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp 178
Db	422	CTCCTCTCTGTCACTGGTTCGCCCTTGAAGATTGCCATACATATACATATGCAACAACATGG 481
Qy	179	ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys 198
Db	482	ATTATGGGNAAGCTCTTTGTAATGTGCTATTGGCTTTTCTATGGCAACATGTACTGT 541
Qy	199	SerIleLeuLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe 218
Db	542	TCCATTCTCTTCATGACCTGCCTCAGTGTGCAGAGGTATTGGGTTCATCGTGAACCCCATG 601
Qy	219	ThrTyrArgGlyLeuProLysHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla 238
Db	602	GGGCAC---TCCAGGAAGAACGAAACATGGCCATTGGCATCTCCCTGGCAATATGGCTG 658
Qy	239	ThrValPheLeuTyrMetLeuProPhePheIleLeuLysGlnGluTyrTyrLeuValGln 258
Db	659	CTGATTCTGCTGGTCACCATCCCTTTGTATGTGCTGAAGCAGACCATCTTCATCTCTGCC 718
Qy	259	ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db	719	CTGAACATCAGCACCTGTCATGATGTT-----TTGCCTGAGCAG 757
Qy	279	LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle 292
Db	758	CCTCTGGTGGAGACATGTTCAATTACTCTCTCTGGCCATGGGGTCTTCTGTGTC 817
Qy	293	ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla----- 310
Db	818	CCAGCCTTCTCACACCTCTGCCTATGTCTGATGATCAGATGCTGCATCTCTGCC 877
Qy	311	-----TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu 326
Db	878	ATGGATGAAAATCAGAGAAGAAAGAGAGGCCATCAAACATCATTTGTCAGTGTCTG 937
Qy	327	ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisHisAlaAsn 346
Db	938	GCCATGTACTGATCGCTTCACTCCTAGTAACCTTCTGCTGTGGTGCAT----- 988
Qy	347	TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu 363
Db	989	TATTTCTGATTAAGACCGAGCCAGCCATGCTATGCCCTGTACATTTAGCCCTC 1048
Qy	364	CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr 383
Db	1049	TGCCTCTCTACCTTAACAGCTGCATCGACCCCCTTGTCTATTACTTTGTTTCCATGAT 1108
Qy	384	---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu 400
Db	1109	TTCAGGGATCATGCAAGAAGCGCTCTCCTTGGCGAAGTGTCCGCATGTGTAAGACAG 1165
RESULT 35		
HSPAR2B		
LOCUS	HSPAR2B	1289 bp DNA linear PRI 21-JUL-1997
DEFINITION	H.sapiens partial gene for proteinase-activated receptor 2 (1289 BP).	
ACCESSION	U49994	
VERSION	249994.1	GI:1008086
KEYWORDS	Homo sapiens.	
SOURCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE
1 (bases 1 to 1289)
AUTHORS
Nystedt, S.
TITLE
Direct Submission
JOURNAL
Submitted (03-JUL-1995) Sverker Nystedt, Division of Neurobiology,
The Wallenberg, Laboratory, Lund University, Sölvegatan 33A, Lund,
S-223 62, Sweden
2 (bases 1 to 1289)
REFERENCE
Nystedt, S., Emilsson, K., Larsson, A.K., Strombeck, B. and Sundelin, J.
AUTHORS
TITLE
Molecular cloning and functional expression of the gene encoding
the human proteinase-activated receptor 2
JOURNAL
Eur. J. Biochem. 232 (1), 84-89 (1995)
MEDLINE
96048032
PUBMED
7556175

FEATURES
source

1. 1289 Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="q13"
c1. .96
97. .>1289
/product="proteinase-activated receptor 2"
/usedin="249993:par2_CDS
1209. .>1289
3'UTR 300 a 320 c 270 g 399 t
ORIGIN

Alignment Scores:

Pred. No.: 8,73e-39 Length: 1289
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 9 Gaps: 10

US-09-208-629F-6 (1-408) x HSPAR2B (1-1289)

QY 80 GlyAlaThrIleThrValLysLysCysProGluSerAlaSerHisLeuHisVal 99
Db 168 GGAGTTACAGTGAACAGCTTTTCTGTCGATGAGTTTCTGCATCT----- 215
QY 100 LysAsnAlaThrMetGlyTyLeuThrSerSerLeuThrLysLeuLeuProAlaIle 119
Db 216 -----GTCCTCAGTGGAAACTGACACTGCTCTCTTCAATGTC 257
QY 120 TyrLeuLeuValPheValGlyValProAlaAsnAlaValThrLeuTrpMetLeuPhe 139
Db 258 TACACAATTGTTTGTGGTGGTTTCCCAAGTAACGGCATGGCCCTGTGGTCTTCTT 317
QY 140 PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp 158
Db 318 TTCGAAGTAAGAAAGAACCCCTGCTGCTGATTACATGGCCCAATCTGGCTTGGCTGAC 377
QY 159 PheLeuPheCysValThrLeuProPheLysIleAlaTyHisLeuAsnGlyAsnAsnTrp 178
Db 378 CTCCTCTCTCATCTGGTCCCTTGAAGATGGCTATCACATACATGCAACACTGG 437
QY 179 ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyGlyAsnMetTyCys 198
Db 438 ATTTATGGGAAGCTCTTTGTAATGCTTATTGGCTTTTCTATGGCAACATGTAAGT 497
QY 199 SerIleLeuLeuAlaCysIleSerIleAsnArgTyLeuAlaIleValHisProphe 218
Db 498 TCCATTCTTCATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
QY 219 ThrTyArgGlyLeuProLysHisThrTyAlaLeuValThrCysGlyLeuValTrpAla 238
Db 558 GGGCAC---TCCAGAGAGAGGCAACATTTGCCATTTGCCATTTCCCTGGCAATGCTG 614
QY 239 ThrValPheLeuTyMetLeuProPheIleLeuLysGlnGluTyTyLeuValGln 258
Db 615 CTGATTCTGCTGGTCCACCTCTTTGTATGTCGTAAGACGACCATCTTCTTCCTGCC 674

QY 259 ProAspIleThrThrCysHisAspValHisAsnThrCysGluSerSerProPheGln 278
Db 675 CTGAACATCAGCACTGTCAATGATGTT-----TTGGCTGAGCAG 713
QY 279 LeuTyR-----TyrPheIleSerLeuAlaPhePheGlyPheLeu 292
Db 714 CTCTTGGTGGGAGACATGTTCAATTTCTCTCTCTGGCCATTTGGGTCTTTCTGTTTC 773
QY 293 PropheValLeuIleTyCysTyAlaAlaIleIleArgThrLeuAsnAla----- 310
Db 774 CCAGCCTTCTCACAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
QY 311 -----TyrAspHisArgTrpLeuTyrValLysAlaSerLeuLeuLeu 326
Db 834 ATGGATGAACACTCAG 893
QY 327 ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn 346
Db 894 GCCATGTACCTGATCTGCTTCACTCTAGTAACCTTCTGCTTGTGTGTCAT----- 944
QY 347 TyrTyTyTyAsnAsnThrAspGly-----LeuTyPheIleTyLeuLeuAlaLeu 363
Db 945 TATTTTCTGATTAAAGAGCCAGGCGGCGAGCCCATGCTATGCCCTGTACATTTAGCCCTC 1004
QY 364 CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyTyPheLeuMetSerLysThr 383
Db 1005 TGCTCTCTACCTTAAACAGCTGCATGACCCCTTGTCTATCTACTTTGTTTCACATGAT 1064
QY 384 ---ArgAsnHisSer---ThrAlaTyLeuThrLys*****AsnAspLeuArgGlu 400
Db 1065 TTCAGGATCATGCAAGAAAGAGCTCTCTTTCGCAAGTCCGCACTGTAAGCAG 1121

RESULT 36

AF4000075 18351 bp DNA linear PRI 18-JAN-2002
LOCUS Homo sapiens coagulation factor II (thrombin) receptor-like 1
DEFINITION (F2RL1) gene, complete cds.

ACCESSION AF4000075
VERSION AF4000075.1 GI:15021772

KEYWORDS
SOURCE

ORGANISM

Homo sapiens.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 18351)

AUTHORS Rieder, M.J., Carrington, D.P., Chung, M.-W., Lee, K.L., Poel, C.L.,

Yi, Q. and Nickerson, D.A.

TITLE Direct Submission

JOURNAL Submitted (17-JUL-2001) Molecular Biotechnology, University of

Washington, 1705 NE Pacific, Seattle, WA 98195, USA

COMMENT To cite this work please use: SeattleSNPs. NHLBI Program for

Genomic Applications, UW-FHCRC, Seattle, WA (URL:

http://pga.mbt.washington.edu)

This sequence consists of 2 contigs. The gap between the contigs

is represented as a run of N.

1 10774: contig of 10774 bp in length

10775 10874: gap of unknown length

10875 18351: contig of 7477 bp in length.

FEATURES

Location/Qualifiers

1. 18351

/organism="Homo sapiens"

/db_xref="taxon:9606"

repeat_region 343..492

/rpt_family="MIR"

/rpt_type="dispersed"

variation 467

/frequency="0.09"

/replace="a"

variation 531

/frequency="0.01"

/replace="g"

variation 678

repeat_region	/frequency="0.01"		
	/replace="c"	variation	/replace="a"
	728. .780		3069
	/rpt_family="L2"		/gene="F2RL1"
	/rpt_type=dispersed		/frequency="0.03"
variation	829		/replace="a"
	/frequency="0.03"	variation	3119
	/replace="a"		/gene="F2RL1"
repeat_region	913. .1028		/frequency="0.32"
	/rpt_family="FLAM C"		/replace="t"
	/rpt_type=dispersed	variation	3151
variation	925		/gene="F2RL1"
	/frequency="0.01"		/frequency="0.01"
	/replace="t"	variation	/replace="g"
	1059. .1214		3263
repeat_region	/rpt_family="L3"		/gene="F2RL1"
	/rpt_type=dispersed		/frequency="0.33"
	1292. .1578	variation	/replace="g"
	/rpt_family="AluSg"		3306
	/rpt_type=dispersed		/gene="F2RL1"
variation	1538		/frequency="0.02"
	/frequency="0.01"	variation	/replace="t"
	/replace="g"		3330
variation	1622		/gene="F2RL1"
	/frequency="0.02"		/frequency="0.01"
	/replace="t"	variation	/replace="c"
	1623		3331
variation	/frequency="0.01"		/gene="F2RL1"
	/replace="g"		/frequency="0.07"
repeat_region	1795. .1926	variation	/replace="t"
	/rpt_family="MIR"		3353
	/rpt_type=dispersed		/gene="F2RL1"
variation	2122		/frequency="0.01"
	/frequency="0.01"	variation	/replace="a"
	/replace="g"		3417
variation	2254		/gene="F2RL1"
	/frequency="0.01"		/frequency="0.01"
	/replace="c"	variation	/replace="c"
	2457		4160
variation	/frequency="0.10"		/gene="F2RL1"
	/replace="g"		/frequency="0.01"
gene	2462. .17176	variation	/replace="c"
	/gene="F2RL1"		4456
mRNA	join(2462. .2687,15958. .17176)		/gene="F2RL1"
	/gene="F2RL1"		/frequency="0.01"
	/product="coagulation factor II (thrombin) receptor-like	variation	/replace=""
	1"		4500
variation	2561		/gene="F2RL1"
	/gene="F2RL1"		/frequency="0.02"
	/frequency="0.28"	variation	/replace="t"
	/replace="c"		4663
	join(2606. .2687,15958. .17069)		/gene="F2RL1"
CDS	/gene="F2RL1"		/frequency="0.46"
	/codon_start=1	repeat_region	/replace="t"
	/product="coagulation factor II (thrombin) receptor-like		4710. .4817
	1"	variation	/rpt_family="L2"
	/protein_id="AAK77914.1"		/rpt_type=dispersed
	/db_xref="GI:15021773"		4715
	/translation="MRSPSAWLLGAILLAASLSCGTTQGTNRSSKGRSLIGKVDG		/gene="F2RL1"
	TSHTGKGVETVFSDEFSASVLTGKLTVPITVITVIVVVGPLSGNALWVLF		/frequency="0.15"
	RTKKHPAVIYMANLADLLSVIFPLKIAYHIGNNWIYGEALCNVLGFFYGNMY	variation	/replace="g"
	CSLFTWCLSVQVWTVNPMGSHRKKANIAIGISLAILLLVITPLVYVVKQTIPI		4972
	PALNITTCBVLPEQLVGMFNYFLSLAIGVFLPFAFLTASAVMLMRLRSSAMDE		/gene="F2RL1"
	NSKKRRATKLVTVLAMYLICFTPSNLLLVVHYFLIKSGQSHVYALVIVALCLST		/frequency="0.02"
	LNSCIDPVVYFVSHDFRDHAKNALLCRSVTVKQVSLTSKHSKSSSSSSSTT	repeat_region	/replace="c"
	VKUSY"		4993. .5299
	2667		/rpt_family="AluSx"
variation	/gene="F2RL1"		/rpt_type=dispersed
	/frequency="0.02"	variation	5065
	/replace="t"		/gene="F2RL1"
	2770		/frequency="0.01"
variation	/gene="F2RL1"		/replace="g"
	/frequency="0.09"	variation	5159
			/gene="F2RL1"

* 1255 2766: contig of 1512 bp in length
* 2767 2866: gap of 100 bp
* 2867 4073: contig of 1207 bp in length
* 4074 4173: gap of 100 bp
* 4174 5191: contig of 1018 bp in length
* 5192 5291: gap of 100 bp
* 5292 7387: contig of 2096 bp in length
* 7388 7487: gap of 100 bp
* 7488 9087: contig of 1600 bp in length
* 9088 9187: gap of 100 bp
* 9188 10361: contig of 1174 bp in length
* 10362 10461: gap of 100 bp
* 10462 12065: contig of 1604 bp in length
* 12066 12165: gap of 100 bp
* 12166 13498: contig of 1333 bp in length
* 13499 13598: gap of 100 bp
* 13599 15085: contig of 1487 bp in length
* 15086 15185: gap of 100 bp
* 15186 17249: contig of 2064 bp in length
* 17250 17349: gap of 100 bp
* 17350 19243: contig of 1894 bp in length
* 19244 19343: gap of 100 bp
* 19344 21127: contig of 1784 bp in length
* 21128 21227: gap of 100 bp
* 21228 23345: contig of 2118 bp in length
* 23346 23445: gap of 100 bp
* 23446 25965: contig of 2520 bp in length
* 25966 26065: gap of 100 bp
* 26066 28195: contig of 2130 bp in length
* 28196 28295: gap of 100 bp
* 28296 32216: contig of 3921 bp in length
* 32217 32316: gap of 100 bp
* 32317 36954: contig of 4638 bp in length
* 36955 37054: gap of 100 bp
* 37055 40348: contig of 3294 bp in length
* 40349 40448: gap of 100 bp
* 40449 43556: contig of 3108 bp in length
* 43557 43656: gap of 100 bp
* 43657 47926: contig of 4270 bp in length
* 47927 48026: gap of 100 bp
* 48027 54079: contig of 6053 bp in length
* 54080 54179: gap of 100 bp
* 54180 58962: contig of 4783 bp in length
* 58963 59062: gap of 100 bp
* 59063 63739: contig of 4677 bp in length
* 63740 63839: gap of 100 bp
* 63840 71551: contig of 7712 bp in length
* 71552 71651: gap of 100 bp
* 71652 78218: contig of 6567 bp in length
* 78219 78318: gap of 100 bp
* 78319 84982: contig of 6664 bp in length
* 84983 85082: gap of 100 bp
* 85083 91930: contig of 6848 bp in length
* 91931 92030: gap of 100 bp
* 92031 98922: contig of 6892 bp in length
* 98923 99022: gap of 100 bp
* 99023 106495: contig of 7474 bp in length
* 106497 106596: gap of 100 bp
* 106597 114950: contig of 8354 bp in length
* 114951 115050: gap of 100 bp
* 115051 124353: contig of 9303 bp in length
* 124354 124453: gap of 100 bp
* 124454 132639: contig of 8186 bp in length
* 132640 132739: gap of 100 bp
* 132740 141572: contig of 8833 bp in length
* 141573 141672: gap of 100 bp
* 141673 151302: contig of 9630 bp in length
* 151303 151402: gap of 100 bp
* 151403 161767: contig of 10365 bp in length
* 161768 161867: gap of 100 bp
* 161868 184536: contig of 22669 bp in length.
Location/Qualifiers
1. 184536

misc_feature
1..1154
/note="assembly_fragment"
vector_end:SP6
vector_side:left
1255..2766
/note="assembly_fragment"
2867..4073
/note="assembly_fragment"
4174..5191
/note="assembly_fragment"
5292..7387
/note="assembly_fragment"
7488..9087
/note="assembly_fragment"
9188..10361
/note="assembly_fragment"
10462..12065
/note="assembly_fragment"
12166..13498
/note="assembly_fragment"
13599..15085
/note="assembly_fragment"
15186..17249
/note="assembly_fragment"
17350..19243
/note="assembly_fragment"
19344..21127
/note="assembly_fragment"
21228..23345
/note="assembly_fragment"
23446..25965
/note="assembly_fragment"
26066..28195
/note="assembly_fragment"
28296..32216
/note="assembly_fragment"
32317..36954
/note="assembly_fragment"
37055..40348
/note="assembly_fragment"
40449..43556
/note="assembly_fragment"
43657..47926
/note="assembly_fragment"
48027..54079
/note="assembly_fragment"
54180..58962
/note="assembly_fragment"
59063..63739
/note="assembly_fragment"

Alignment Scores:
Pred. No.: 2 62e-36 Length: 184536
Score: 585.50 Matches: 124
Percent Similarity: 54.87% Conservative: 62
Best Local Similarity: 36.58% Mismatches: 114
Query Match: 27.14% Indels: 39
DB: 2 Gaps: 10

US-09-208-629f-6 (1-408) x AC068682 (1-184536)

Qy 80 GlyAlaThrIleThrValLysIleLysCysProGluGluSerAlaSerHisLeuHisVal 99
Db 66691 GGAGTTACAGTTCGAACAGCTCTTTCTGTGATGAGCTTTTCTGCATCT----- 66644
Qy 100 LysAsnAlaThrMetGlyTyrLeuThrSerLeuSerLeuSerLeuProAlaIle 119
Db 66643 -----GTCCTCTGGAACACTGACCACCTCTCTCTCAATTGTC 66602

Qy	120	TyrLeuLeuValPheValValGlyValProAlaAsnAlaValThrLeuTrpMeLeuPhe	139
Db	66601	TACACAATTGTGTGTTGTGGTGGTTTGGCAAGTAACGCATGGCCCTCTGCTTCTTTCTT	66542
Qy	140	PheArgThrArgSerIleCysThrThrValPheTyr---ThrAsnLeuAlaIleAlaAsp	158
Db	66541	TTCCGNACTTAAAGAGAGACCCTGCTGTGATTATACATGGCCAATCTGGCCCTTGCGTGAC	66482
Qy	159	PheLeuPheCysValThrLeuPhePheLysIleAlaTyrHisLeuAsnGlyAsnAsnTrp	178
Db	66481	CTCCTCTCTGTATCTGGTTCCCTTGAAGATTGCTATCATACATACATGGCAACAACCTGG	66422
Qy	179	ValPheGlyGluValLeuCysArgAlaThrThrValIlePheTyrGlyAsnMetTyrCys	198
Db	66421	ATTATTGGGGAAGCTCTTTGTAACTGTGCTTATTGGCTTTTCTATGGCAACATGTACTGT	66362
Qy	199	SerIleLeuLeuAlaCysIleSerIleAsnArgTyrLeuAlaIleValHisProPhe	218
Db	66361	TCAATTCTCTCATGACCTGCTCATGTGCGAGAGTATGGGTATCTGTAACCCCATG	66302
Qy	219	ThrTyrArgGlyLeuProIlyHisThrTyrAlaLeuValThrCysGlyLeuValTrpAla	238
Db	66301	GGGCAC---TCCAGGAAGAGCAACATTGGCATCTCCCTGGCAATATGCTG	66245
Qy	239	ThrValPheLeuTyrMetLeuPhePheIleLeuLysGlnGluTyrTyrLeuValGln	258
Db	66244	CTGATTCTGCTGGTCACCATCCCTTTGATGTCGTGAAGCAGACCATCTTCATTCCTGCC	66185
Qy	259	ProAspIleThrThrCysHisAspValHisThrCysGluSerSerProPheGln	278
Db	66184	CTGAACATCAGCACCTGTCATGTT-----TTCCCTTGCAGCAG	66146
Qy	279	LeuTyr-----TyrPheIleSerLeuAlaPhePheGlyPheLeuIle	292
Db	66145	CTCTTGGTGGGAGACATGTTCAATTACTTCTCTCTGGCCATGGGCTCTTCTGTTC	66086
Qy	293	ProPheValLeuIleIleTyrCysTyrAlaAlaIleIleArgThrLeuAsnAla-----	310
Db	66085	CCAGCCTTCTCACAGCCTCTGCCTATGTGCTGATGATCAGATGCTCGCATCTTCTGCC	66026
Qy	311	-----TyrAspHisArgTrpLeuTrpTyrValLysAlaSerLeuLeuIleLeu	326
Db	66025	ATGGATGAAACTCAGAGAGAGAGAGAGAGGCGCCATCAAACTCATGTCACTGTCTGTG	65966
Qy	327	ValIlePheThrIleCysPheAlaProSerAsnIleIleLeuIleIleHisAlaAsn	346
Db	65965	GCCATGTACTGATCTGCTTCACTCTAGTAACCTTCTGCTTGTGGTGAT-----	65915
Qy	347	TyrTyrTyrAsnAsnThrAspGly-----LeuTyrPheIleTyrLeuIleAlaLeu	363
Db	65914	TATTTTCTGATTAAAGAGCCAGGCGCAGAGCCATGCTATGCGCTGTACATGTAGCCCTC	65855
Qy	364	CysLeuGlySerLeuAsnSerCysLeuAspProPheLeuTyrPheLeuMetSerLysThr	383
Db	65854	TGCCTCTCAACCTTAAACAGCTGCATCGACCCCTTTGTCTATATTCTTTTCATCAT	65795
Qy	384	---ArgAsnHisSer---ThrAlaTyrLeuThrLys*****AsnAspLeuArgGlu	400
Db	65794	TTCCGGGATCATGCAAGAACGCTCTCCTTTGCCGAAGATGTCGACATGTAAGACG	65738
RESULT	39		
LOCUS	AC114962		
DEFINITION	Homo sapiens chromosome 5 clone RP11-206N2, WORKING DRAFT SEQUENCE,	DNA	linear
	8 unordered pieces.		
ACCESSION	AC114962		
VERSION	AC114962.1	GI:19424423	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 20485)		

AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT

```
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 204485)
DOE Joint Genome Institute.
Direct Submission
Submitted (14-MAR-2002) production
Genome Institute, 2800 Mitchell
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 485783
Center clone name: RPCI-11.206N2
```

Summary Statistics

Consensus quality:	195215 bases at least Q40
Consensus quality:	199004 bases at least Q30
Consensus quality:	200415 bases at least Q20
Estimated insert size:	175000; agarose-fp estimation
Estimated insert size:	203785; sum-of-contigs estimation
Quality coverage:	9.78 in Q20 bases; agarose-fp estimation
Quality coverage:	8.4 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently

- * consists of 8 contigs. The true order of the pieces
- * is not known and their order in this sequence record is
- * arbitrary. Gaps between the contigs are represented as
- * runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will
- * be preserved.

	1	1058:	contig of 1058 bp in length
*	1059	1158:	gap of unknown length
*	1159	2342:	contig of 1184 bp in length
*	2343	2442:	gap of unknown length
*	2443	4307:	contig of 2465 bp in length
*	4908	5007:	gap of unknown length
*	5008	19885:	contig of 14878 bp in length
*	19886	19985:	gap of unknown length
*	19986	35850:	contig of 15865 bp in length
*	35851	35950:	gap of unknown length
*	35951	53802:	contig of 17852 bp in length
*	53803	53902:	gap of unknown length
*	53903	107788:	contig of 53886 bp in length
*	107789	107888:	gap of unknown length
*	107889	204485:	contig of 96597 bp in length

FEATURES
source

```

1 204485 /organism="Homo sapiens"
/ab_xref="taxon:9606"
/chromosome="5"
/clone="RP11-206N2"
/clone_lib="RP11 human B"
56752 a 44584 c 44954 g 5749
BASE COUNT
ORIGIN

```

Alignment Scores:

```

Assignment Scores:
Pred. No.:      2.95e-36
Score:          585.50
Percent Similarity: 54.87%
Best Local Similarity: 36.58%
Query Match:    27.14%
DB:             2

```

US-09-208-629F-6 (1-408) X AC114962 (1-204485)

Qy 80 GlyAlaThrIleThrValLysileLysCysProGluGlusSerAlaSerHisLeuHisVal 99
||| |||::: ||| ||| ||| ||| |||
Db 73897 GGAGTTACAGTGTAAACACAGCTCTTCTTGATGATGAGTTCTTCGCATCT----- 73944

Qy 100 LysAsnAlaThrMetGlyTyrlLeuthrSerSerLeuSerThrLysLeuileproAlaile 119


```
Db 734 GTT-----TTGCCCTGAGCAGCTCTTGGTGGGAGACATGTTCAAT 772
Qy 281 TyrPheIleSerLeuAlaPhePheGlyPheLeuIleProPheValIleIleIleTyrCys 300
Db 773 TACTTCCTCTCTCTGGCCATGGGGCTTTCTGTTCCTCCAGCCTTCTCAGAGCCTCTGCC 832
Qy 301 TyrAlaAlaIleIleArgThrLeuAsnAla-----TyrAspHisArg 314
Db 833 TATGTGCTGATGATCAGAAATGCTGCGATCTTCTGCCATGGATGAAAACTCAGAGAGAAA 892
Qy 315 TrpLeuTrpTyrValIysAlaSerLeuLeuIleValIlePheThrIleCysPheAla 334
Db 893 AGGAAGAGGGCCATCAAACTCATTTGCTCTGGGCAATGACCTGTGCTTCACT 952
Qy 335 ProSerAsnIleIleLeuIleHisAlaAsnTyrTyrTyrAsnAsnThrAspGly 354
Db 953 CCTAGTAACCTTCTGCTTGTGGTGCAT-----TATTTCTGATTAGAGCCAGGCG 1003
Qy 355 -----LeuTyrPheIleTyrLeuIleAlaLeuCysLeuGlySerLeuAsnSerCys 371
Db 1004 CAGAGCCATGCTATGCGCTGTACATGTAGCCCTCTGCCCTCTTACCCCTTAACAGCTGC 1063
Qy 372 LeuAspProPheLeuTyrPheLeuMetSerLysThr---ArgAsnHisSer---ThrAla 389
Db 1064 ATCGACCCCTTGTCTATTACTTTGTTTCACATGATTTACGGGATCATGCAAGAAGCGCT 1123
Qy 390 TyrLeuThrLys*****AsnAspLeuArgGluGlnGlyGlnPro 404
Db 1124 CTCCTTTGCCGAAGTGTCCGCACTGTAAAGCAGATGCAAGTACCC 1169
```

Search completed: June 29, 2003, 09:48:53
Job time : 2761.78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:58:31 ; Search time 51.5632 Seconds
(without alignments)
1630.373 Million cell updates/sec

Title: US-09-208-629f-6
Perfect score: 2157
Sequence: 1 CSMLQISXLRDGTQVIM.....AYLTXXNDLRQEQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues 671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rvirus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	605.5	28.1	399	11 Q8R311	Q8R311 mus musculus
2	453	21.0	357	13 Q9D805	Q9D805 raja erinac
3	432.5	20.1	374	13 Q57466	Q57466 meleagris g
4	422.5	19.6	361	13 Q90X57	Q90X57 xenopus lae
5	387	17.9	359	11 Q99MT7	Q99MT7 mus musculus
6	368.5	17.1	358	4 Q9BY21	Q9BY21 homo sapien
7	368.5	17.1	358	4 Q96J28	Q96J28 homo sapien
8	367.5	17.0	309	11 Q8R528	Q8R528 mus musculus
9	367.5	17.0	359	13 Q9PVY7	Q9PVY7 anguilla an
10	358.5	16.6	342	6 Q9GK76	Q9GK76 capra hircu
11	356.5	16.5	342	6 Q9TTY5	Q9TTY5 bos taurus
12	353	16.4	367	11 Q9RLV0	Q9RLV0 mus musculus
13	351	16.3	337	4 Q96P68	Q96P68 homo sapien
14	347	16.1	361	6 Q46685	Q46685 bos taurus
15	345	16.0	359	11 Q9EPP3	Q9EPP3 cavia porce
16	341.5	15.8	296	6 Q9TTY6	Q9TTY6 canis famil

17	340	15.8	328	11 Q9ERK9	Q9ERK9 mus musculu
18	340	15.8	359	6 Q9NOU1	Q9NOU1 ovis aries
19	338	15.7	334	6 Q97724	Q97724 felis silve
20	337.5	15.6	298	11 Q8VE54	Q8VE54 mus musculu
21	337	15.6	300	13 Q8QGO4	Q8QGO4 carassius a
22	336.5	15.6	342	6 Q9XSD4	Q9XSD4 sus scrofa
23	334	15.5	359	6 Q9GLN9	Q9GLN9 pan troglod
24	333.5	15.5	339	6 Q9TUS6	Q9TUS6 papio cynoc
25	333	15.4	353	13 Q8QFR5	Q8QFR5 xenopus lae
26	332	15.4	380	4 Q9UNW8	Q9UNW8 homo sapien
27	331.5	15.4	352	6 Q9TV48	Q9TV48 cercopithec
28	331.5	15.4	382	11 Q9Z282	Q9Z282 mus musculu
29	331	15.3	359	4 Q8TBK4	Q8TBK4 homo sapien
30	330.5	15.3	339	6 Q9TUR6	Q9TUR6 cercopithec
31	330	15.3	342	13 Q93239	Q93239 cyprinus ca
32	329.5	15.3	339	6 Q9TUU5	Q9TUU5 macaca fusc
33	329.5	15.3	339	6 Q9TUR4	Q9TUR4 mandrillus
34	329.5	15.3	352	6 Q95NC3	Q95NC3 miopithecus
35	328.5	15.2	339	6 Q9TUU8	Q9TUU8 cercopithec
36	328.5	15.2	339	6 Q9TSQ3	Q9TSQ3 cercopithec
37	328.5	15.2	343	6 Q9NOZ0	Q9NOZ0 cercocebus
38	328	15.2	355	6 Q9MYJ8	Q9MYJ8 callithrix
39	327.5	15.2	339	6 Q9TUU6	Q9TUU6 macaca fusc
40	327.5	15.2	339	6 Q9TQX3	Q9TQX3 mandrillus
41	327.5	15.2	352	6 Q95ND2	Q95ND2 mandrillus
42	327	15.2	337	11 Q921N3	Q921N3 mus musculu
43	326.5	15.1	343	6 Q9BDS6	Q9BDS6 macaca fusc
44	325.5	15.1	339	6 Q9TUQ8	Q9TUQ8 cercopithec
45	325.5	15.1	342	6 Q9TV16	Q9TV16 pan troglod

ALIGNMENTS

RESULT 1

Q8R311 ID Q8R311 PRELIMINARY; PRT; 399 AA.
AC Q8R311;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor II (thrombin) receptor-like 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025432; AAH25432.1; -;
KW Receptor.
SQ SEQUENCE 399 AA; 44679 MW; 807C79464AB9B3EF CRC64;

Query Match	28.1%	Score	605.5;	DB 11;	Length	399;
Best Local Similarity	35.4%;	Pred. No.	4.2e-44;			
Matches	137;	Conservative	68;	Mismatches	127;	Indels 55; Gaps 12;
QY	22	ALIFAAAGLLLLPTFCQSGMEN--DTNNLAKPTLPKTRGAPPSPFEFFPSALEGWT	79			
Db	5	SLAWLGGITLLAASVCSRTENLAPGRNNKSGRSLGRLETOPP-----IT	51			
QY	80	GATITVK--IKCPESASHLHVKNATMGVLTSSLTSLKLIPIAYLLVFGVGPANAVTLWM	137			
Db	52	GKGVPVEPGSIDFAS-----ILTKLTTFVFLPVVIIIVFVIGLPSNGMALWI	101			
QY	138	LFRTSICITTFY--TNLAIADELFVCLPFIKAYHLNGNNWVGEVCLCRATTVFYGNM	196			
Db	102	FLFRTKKHPAIVYMANLADLLSVTFPLFAYHLHGNWVYVGEALCKVLIGFFYGNM	161			
QY	197	YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCLG---VWATVFLVWLPFFILKQE	253			
Db	162	YCSILFMTCLSVQRYWYVIVNPM---GHPKKK-ANIAVGSVLAIWLLIFLVTIPLYVMKOT	217			

```
Qy 254 YLVQPDITTCVHNTCESSPPQLY-----YFISLAFGFLIPFLVLIYCVAAIIRT 307
Db 218 IYIPALNITTCADV-----LPEVLVGMFNFLSLAIGVFLFPAILTASAYVLMK 270
Qy 308 LNA-----YDHRWLWYKASLLILVITTCFAPSNIILIHANYNNYNTDGLFYILI 361
Db 271 LRSSAMDEBSKKQRAIRLITITVLAMVYFCFAPSNNLLVVHYFLINKTQRQSHVYALV 330
Qy 362 ALCLGSLNSCLDPLFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYVYFVSKDFRDA 357

RESULT 2
Q9DE05 PRELIMINARY; PRT; 357 AA.
ID AC Q9DE05
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE P2Y receptor.
OS Raja erinacea (little skate).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalae; Hymnosqualea; Pristiogalea; Batoidae;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=7782;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20459151; PubMed=10900200;
RA Dranoff J.A., O'Neill A.F., Franco A.M., Cai S.Y., Connolly G.C.,
RA Ballatori N., Boyer J.L., Nathanson M.H.;
RT "A primitive ATP receptor from the little skate Raja erinacea.";
RL J. Biol. Chem. 275:30701-30706 (2000).
DR EMBL; AF242850; AAG42684.1; -
DR HSSP; P34996; 1DDD.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR RECEPTOR.
KW RECEPTOR.
SQ SEQUENCE 357 AA; 41239 MW; 14604E15DCBDB41 CRC64;

Query Match 21.0%; Score 453; DB 13; Length 357;
Best Local Similarity 32.9%; Pred. No. 4.8e-31;
Matches 102; Conservative 54; Mismatches 132; Indels 22; Gaps 8;

Qy 116 IPAILLVVGVGPANAVTLWMLFFRTSIC-TVFVTNLAIAADFLFCVTLPPKIAVHLN 174
Db 38 LPIMVIVVFTGTSVALWMLFPHRPWSIIHYMFLNVLADLFVFSLPILIFYFN 97
Qy 175 GNNVFGVLCRATTVPYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCG 234
Db 98 KTDWIFGELLCKLRFIFHVLNYSILFLTCSVHYRTGVVHPMKSLGRLKKSATIVCV 157
Qy 235 LWATVFLYMLPFFILKQEYLVQPDITTCVHNTCESSPPQLYFISLAFGLIIP 294
Db 158 CWITWAGISPIIFYST-GLRRNKNTCYD--TTSKELLETFYISMSTTFPGFCIPF 214
Qy 295 VLIIVCYAAIIRTNAYDHRWLW---YVKASLLILVITTCFAPSNIILIH-HANYYN 350
Db 215 ATILVCYGFIVKALISNDMKTPLRGKSVRLVIVLAVFAISYLFPHWKNLNLOSRIYQ 274
Qy 351 NTD-----GLYFYIYIALCLGSLNSCLDPLFLMSKT-----RNHSTAYLTXXN--- 396
Db 275 GLDTCWNNRRVYATVQVTRGLASLNSCVDPILYFLAGDTFRRRTTNAASRPMTRSEL 334
Qy 397 DLREQQPSQ 406
Db 335 QFRSEDSPLQ 344
```

```
RESULT 3
Q57466 PRELIMINARY; PRT; 374 AA.
ID AC Q57466
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE G protein coupled P2Y nucleotide receptor.
GN TP2Y.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
OX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE=98086419; PubMed=9415702;
RA Boyer J.L., Waldo G.L., Harden T.K.;
RT "Molecular cloning and expression of an avian G protein-coupled P2Y
RT receptor.";
RL Mol. Pharmacol. 52:928-934 (1997).
DR EMBL; AF031897; AAC60339.1; -
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR RECEPTOR.
KW RECEPTOR.
SQ SEQUENCE 374 AA; 42594 MW; 849C465722BDD02B CRC64;

Query Match 20.1%; Score 432.5; DB 13; Length 374;
Best Local Similarity 30.0%; Pred. No. 2.9e-29;
Matches 106; Conservative 57; Mismatches 123; Indels 67; Gaps 11;

Qy 55 PIKTRGAPPNSPEFPFSALEGWT-----GATITVTKIKCPESASHLHVKNATMG 105
Db 4 FVRMFLAP-----WTPTPWLGGNTAAAEAKC-----VFNEEFK 40
Qy 106 YLTSSSLTKLIPAILLVVGVGPANAVTLWMLFFRTS-ICTVFTYNLAIAADFLFCVT 164
Db 41 FL-----LLPISYGVIVVGVGLPNSWAMWIFVSRMRPNATTYMFNLAISDTLYVFS 93
Qy 165 LPFKIAVHLNGNNWVGEVLCRATTVFYGNMYCSILLACISINRYLAIVHPFTYRGLP 224
Db 94 LPLVTYYADRNWNPFGKVFCKIVRFLFYANLYSSILFLTCISVHRFMGICHPIRSLKWV 153
Qy 225 KHTYALVTCGLVWATVFLYMLP--FFILKQEYLVQPDITTCVHNTCESSPPQLYF 282
Db 154 KTKHARLICVGVWLVVITICLIPNLIFVTSS-----KDNSTLCHD--TTKPEEFDHYVHS 207
Qy 283 ISLAFFGLIPFLVLIYCVAAIIRTL-----NAYDHRWLWYKASLLILVIFT 330
Db 208 SSMALLFGIPFLVIVVVCYCLMAKRLCKRSFPSPRPVSKYKR---SIKMIIVLTVFA 264
Qy 331 ICFAPSNILIIHHANYVY---NNTDGLYFIYIALCLGSLNSCLDPLFLYEL 379
Db 265 ICFVPHITRTLYTSRYFOADCOTLNIINFTYKTRPLASINSLDPLIFM 317

RESULT 4
Q90X57 PRELIMINARY; PRT; 361 AA.
ID AC Q90X57
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE P2Y nucleotide receptor.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
```

```
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheng A.W., Tsing K.W.;
RT "Cloning of Xenopus P2y1 Receptor.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF432354; AAL27614.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 361 AA; 41002 MW; E5B2D605F5B57FED CRC64;

Query Match 19.6%; Score 422.5; DB 13; Length 361;
Best Local Similarity 31.6%; Pred. No. 2e-28;
Matches 107; Conservative 56; Mismatches 137; Indels 39; Gaps 9;

QY 95 SHLVKATMGYLTSLSLK-----LIPALYLLVFGVGPANAVTLMLFFRTRISCT- 147
Db 15 STLASGSSAGNVTKSLTKTGQFYPLPAVYIVCITGFIGNSVAIWMFIFHKPWSI 74
QY 148 TVFTNLAIADFLFCVTLPPKIAVHLNGNNVFEVLCRAITVIFYGNMYCSILLACIS 207
Db 75 SVYFNALADFLVLSLPAIFVFNKTDWIFGALCKLQRFHLVNLVGSILFLTCIS 134
QY 208 INRYLAIVHPPTYRGLPKHTYALVTCGLVWATVFLYMLP--FF----ILKQEYVLPDI 261
Db 135 VHYTGVVHPLKSLGRUKKNSIYISALVWPIVAGISPILPFGSGTGIRKNK-----T 187
QY 262 TTHCHVNTCESSPPQLYYFI----SLAFGLIPFPVLIYCYAAIIRTLNAYDHRWL 318
Db 188 ITCEDT-----SSDEYLSRYFIYSMCTTVFGFCIPFILILGCVLIYRALIYKDMNAPL 242
QY 319 VKAS---LLILVITFCFAPSNIILIH-----HANYVYNTDGLYFLYILALCLGS 367
Db 243 RKXSIYLVIVLTFAVSYLFPFWKKNLRLARLDFOSPENCNDRVYATYQVTRGLAS 302
QY 368 LNSCLDPFLYFLMSKTRNHSTAYLTQXKNDLREQQPSQ 406
Db 303 LNSCVDPILYFLAGDTFRKLSRATRKASRRSEANVQSK 341

RESULT 5
Q99MT7 PRELIMINARY; PRT; 359 AA.
AC Q99MT7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE "An expressed sequence tag (est) data mining strategy succeeding in the discovery of new g-protein coupled receptors.";
GN GPR87.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AF295366; AAK01866.1; -.
DR MGI: 1934133; Gpr87.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41462 MW; 2EB313C273BFCD6A CRC64;

Query Match 17.9%; Score 387; DB 11; Length 359;
Best Local Similarity 28.2%; Pred. No. 2.2e-25;
Matches 100; Conservative 70; Mismatches 131; Indels 54; Gaps 11;

QY 49 LAKPTLPKTRGAPPNSFEFFPSALEGWGTATITVKIKCPESASHLVKATMGYL 108
Db 1 MAVENNVNSTFAGN-----ELYSQASHTA-----NSTSEGHGKNSL--LH 38
QY 109 SSLSTKLIPALYLLVFGVGPANAVTLMLFFRTRISCTTVFY-TNLAIADFLFCVTL 167
Db 39 NKEFTILPVLVIVFVASILLNGLAW-IFPHIRNKTSFIYLNKIVVADLMTLTFPF 97
QY 168 KIAYHLNGNNVFEVLCRAITVIFYGNMYCSILLACISINRYLAIVHPPTYRGLPKHT 227
Db 98 RIVRDAGGFWYEFILCRYTSVLFYANMYTSIVFLGLISVDRLYLVKVPFGDSRMYSIT 157
QY 228 YALVTCGLVWATVFLYMLPFFILKQEYVLPDIITCHDVHNTCESSPPQLYYFISLAF 287
Db 158 FTKVLSVCVWVIMAILSLPNILTNG---QP---TKENIHDCWKLKSLGCAKWHMAVY 210
QY 288 FG---FLIPVLIYCYAAIIRTLNAYDHRWL-----WYVKASLLILVITFCFAPS 337
Db 211 VDSCLFVAVLVILIGCVIAISRYIHKSSROFISQSRKRKHNSIRVVAVFTCFPLPH 270
QY 338 IILIHANHYVYNTDQ-----LYFYILALCLGSLNSCLDPFLYFLMSKT 383
Db 271 LCRI-----PTFSNLRDLDESAAHKILYCKEWTFLSACNVCLDPIIYFMCKS 321

RESULT 6
Q9BY21 PRELIMINARY; PRT; 358 AA.
AC Q9BY21;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE "An expressed sequence tag (est) data mining strategy succeeding in the discovery of new g-protein coupled receptors.";
GN GPR87 OR GPR95.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=21172992; PubMed=11273702;
RA Wittenberger T., Schaller H.C., Hellebrand S.;
RT "An expressed sequence tag (est) data mining strategy succeeding in the discovery of new g-protein coupled receptors.";
RL J. Mol. Biol. 307:799-813(2001).
[2]
SEQUENCE FROM N.A.
MEDLINE=21458557; PubMed=11574155;
RA Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,
RA Lewis T., Evans J.F., George S.R., O'Dowd B.F.;
RT "Discovery and mapping of ten novel G protein-coupled receptor genes.";
RL Gene 275:83-91(2001).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
DR EMBL: AF237763; AAK01858.1; -.
DR EMBL: AF411114; AAL26485.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 358 AA; 41435 MW; FDF157295B5E5D10F CRC64;
```

Query Match	17.1%;	Score 368.5;	DB 4;	Length 358;
Best Local Similarity	29.3%;	Pred. No. 8.6e-24;		
Matches	89;	Conservative 64;	Mismatches 116;	Indels 35; Gaps 9;
QY	100	KNATMGYLTSLSLTKLIPAI	LVLLVVGVPANAVTLM	LFFPFRSICITVFY-TNLAT 158
DB	32	KNTT---	LHNEPDTVLVPLVLIIF	VASILLGLAW-IPFHIRNTSIFLYLKNIVAD 87
QY	159	FLFCVTLPPFKIAYHLNGN	WVGEVLCRATTVIPYGN	MGCSILLLACISINRYLAIVHPF 218
DB	88	LIMTLTFPFRIVHDAGFG	PPWYFKELCRYLSVL	FYANMYTSIVFLGLISIDRYLKVVRPF 147
QY	219	TYRGLPKHTYALVTCGL	WAVTVLYMLPFFIL	KOEYVLVQPDITTCVDVHNTCESSSPFQ 278
DB	148	GDGRMYSITFTKVL	SVCVVMVMAVLSLEN	IITLNG---QP---TEDNTHDCSKLSPGL 200
QY	279	LYYFSLAPFG---	FLIPFLVLIICYAAI	RTLNAYDHRWL-----WVVKASLLILVI 328
DB	201	VKMHTAVTVNSCLFV	AVLVILIGYIAISRY	HKSSRQFISOSSRRKRKHNSIRVVAV 260
QY	329	FRICFAPSNIILIIH	HANYNNYNTDG-----	LYFIYLIACILGSLNSCLDPFLYPL 379
DB	261	FTFCFLPYHLCR---	PFTSHLDRLLDESA	KILYYCKEITLFLSACNVCLDPIYFF 316
QY	380	MSKT 383		
DB	317	MCRS 320		

```

RESULT 7
Q96JZ8 PRELIMINARY; PRT; 358 AA.
ID AC Q96JZ8;
DT 01-DEC-2001 (trEMBLrel. 19, Created)
DT 01-DEC-2001 (trEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (trEMBLrel. 21, Last annotation update)
DE CDNA FLU14878 fis, clone PLACE1003238, weakly similar to probable G
DE protein-coupled receptor.
DE OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI TaxID=9606;
[1]
RN
RP SEQUENCE FROM N.A.
RC
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuna M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Iehi S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Negahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBDJ databases.
RL EMBL; AK027784; BAB55366.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01569; P2Y12PRNCPTR.
DR PRINTS; PR01655; UDPGLUCOSER.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS02462; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS02462; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 358 AA; 41462 MW; 7632700c8e3cc7b9 CRC64;

```

Query Match	17.1%	Score 368.5;	DB 4;	Length 358;
Best Local Similarity	29.3%;	Pred. No. 8.6e-24;		
Matches	89;	Conservative 64;	Mismatches 116;	Indels 35;
Gaps	9			
Qy	100	KNATMGYLTSSLSKLPALVLLVYVVGVPANAVTLAMLFRTSRCTTFFV-TNLAIAD	158	
Db	32	KNTT---LHNEPDTLPVLVLIIIVASILLNGAVW-IFFHRTKTSIFLYLKNIVWAD	87	
Qy	159	FLFCVTLPPKIAHYLNGNNWVGEYLCRATTVIFYGNMYSILLACISINRYLAIVHPF	218	

```

Db 88 LIMLTFFPRFVHDAGEGPWFYFELICRTSVLFPYANMYTSIVFLGLISIDRLVKVKKPF 147
Qy 219 TYRGLPKHTYALVTCGLWATVELYMLPFFILKOEYVLVQPDITTCHDVHNTCESSPFQ 278
Db 148 GDSRMYNITFTKVLSCVVMVIMAVLSLPLNIILTNQ---Qp---TEDNIHDCSKLKSPLG 200
Qy 279 LYFVFISLAFG---FLIPFVLIICYAAIIRTLNAYDHRWL-----WYWKASLLILVI 328
Db 201 VKWHTAVTVNNSCLFVAVLVILGCVYIAISRYTHKSQRQFISOSSRRKRGINQSVVVAV 260
Qy 329 FTTCFAPSNIILIIHHANYNYNTDG-----LYFIYLIACLGSLNSCLDPFLYFL 379
Db 261 FFFCFPLPYHLCRT---PFTFSLDRLLDSEAKILYCKEITLFLSACNVCLDPIIYFF 316
Qy 380 MSKT 383
Db 317 MCRS 320

```

RESULT 8

Q8R528	PRELIMINARY;	PRT;	309 AA.
ID	Q8R528		
AC	Q8R528;		
DT	01-JUN-2002 (TrEMBLrel. 21, Created)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Cysteinyln leukotriene 2 receptor.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RA	Ishii S.;		
RT	"Mouse CysLT2 Gene. ";		
RL	Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.		
EMBL	AB058930; BAB86881.1; -.		
KW	Receptor.		
SO	SEQUENCE.		
SW	SEQUENCE.		
SW	309 AA; 35226 MW; 96FACCB8AF96974 CRC64;		

Query Match 17.0%; Score 367.5; DB 11; Length 309;
Best Local Similarity 30.3%; Pred. No. 9.2e-24;
Matches 89; Conservative 60; Mismatches 114; Indels 31; Gaps 10;

Qy	100	KNATMGVLTSSLSKTLPIALYILVFGVGPANAVTLMMLPFT--RSICTTVFTVNLALIA	157
Db	13	RNCTI-----BNFKKEFYPIIILYIFFMGALNGFSIY-VFLQTCCKSTSVNVFNLNLS	67
Qy	158	DFLFCVTLTPFKIAYHLNGNNWFEVLCRATTVIFYGNMYCSILLLACISINRYLAVHP	217
Db	68	DFLEISLTLPRADYYFPGSNWIFGDLACRVMSYSLYNNMYSIYFLTVLSVRELATVHP	127
Qy	218	FTVRGLPKHYVALVTCGLVWATVFLYMLPFILKQEYVLVQPDITTCCHDVHNTCESSPPF	277
Db	128	FRMPHVTSVRSAIMLCGIIWVFMASSALLLVNGQE---BKDNIISCLEL-----SPQKF	179
Qy	278	Q---LYYFISLAFEGFLIPVLLIYCYAAIIRLTNAYD-----HRLWYWKASLLI	325
Db	180	KSLLIIMNHIAA-VGFLPELFTTICVLLIIRILLKAEIPESGPAAHRA--LTTIVIA	236
Qy	326	LVFTTICFAPSNIILIIHHANYNNNT-DGLYFIYLIALCLGINSCLDPFLYF	378
Db	237	MIYFLLCFLPVHARTLHLVTLWDKSDCVLHKATVITLTMAAANSFENFPLYF	290

RESULT 9
O9PVY7

[illegible]

```
OS Anguilla anguilla (European freshwater eel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
OC Anguilla.
OX NCBI_TaxID=7936;
RN [1]
RP SEQUENCE FROM N.A.
RA Tran Van Chuoi M., Dolphin C.T., Barker S., Clark A.J., Vinson G.P.;
RT "Molecular cloning and characterization of the cDNA encoding the
RT angiotensin II receptor of european eel Anguilla anguilla.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ005132; CAB40835.1; -.
DR HSSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR000130; Zn_MTPeptide.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 39997 MW; 053726AEA827C271 CRC64;

Query Match 17.0%; Score 367.5; DB 13; Length 359;
Best Local Similarity 27.0%; Pred. No. 1.1e-23;
Matches 99; Conservative 61; Mismatches 129; Indels 77; Gaps 10;

QY 99 VKNATMG-----YLTSSLSTK-----LIPAIYLLVFGVGPANAVTLMLF-FRTRSIK 146
DB 1 MENLTGRTGEGIHITCNTSGRHSYIYTLIPVYGCNFGVIGVGNMVAIYCYMKLKV 60
QY 147 TTFTYTNLATDFLCVTLPKIAIYHLNNGNWGEVLCRATTVIFGNMYCSILLACI 206
DB 61 ANIEVLNLAASDLTFLTLPMWATFTAGYGNWPFGLCKASAGLTIFNLYTSIFFLTSL 120
QY 207 SINRYLAIVHPFTVGRGLPKHYALVTCGLVNAVTLVFLPKQEVYLVQPDITTC 266
DB 121 SIDRYLAIVHPVRQRRTVYARITCVLIWAFALLSLPTALSRDFTINHPNTVC-- 178
QY 267 VHNTESSSPQLYFISL--AFFGLIPFLVLIYCYAAIIRTL-----NAYDHRWLWY 318
DB 179 --GTLDKHSLHVLVAIGLMSKVLGFLIPFVIIVTCYCLIGRALLEARRVQSSRSRGEV 236
QY 319 VKASLLILVITFCFAPSNIILITHANYNNYNTDGLYFYLIAM----- 363
DB 237 LQMLAAVVLAAFFLCWVPHQI-----FPHMVLALLKVIENCPTLIDI 280
QY 364 -----CLGSLNSCLDPFLYFLMSKT--RN-----HSTAYLTXXNDLREQ 401
DB 281 ALPTICIAVFNSCWNPILYFGVGNFRNLLRLRLRCGPGSAARHSPSLITKMTLSYR 340
QY 402 GQPSQR 407
DB 341 ASETLR 346

RESULT 10
Q9GK76 PRELIMINARY; PRT; 342 AA.
AC Q9GK76;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RA Yang W., Diehl J.R., Piumi F.;
RT "Partial Sequence of Caprine Platelet-Activating Factor Receptor
RT Gene.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF302764; AAC39982.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCR_RHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 39612 MW; 39D9379AED60FD4E CRC64;

Query Match 16.6%; Score 358.5; DB 6; Length 342;
Best Local Similarity 30.7%; Pred. No. 6e-23;
Matches 99; Conservative 48; Mismatches 129; Indels 47; Gaps 12;

QY 115 LIPAIYLLVFGVGPANAVTLMW---LFFRTRSICTTVFTYNLAIAADFLFCVTLPPKIAY 171
DB 17 LFPFYSIVFVLGVIANSYLVWFAFLYPSKKEIKIFWVNLTMADLLFLVTLPLWIVY 76
QY 172 HLNNGNWGEVLCRATTVIFGNMYCSILLACISINRYLAIVHPFTVGRGLPKHYALV 231
DB 77 YNQGDMILPKFLCNLAGCFFFTNTYCSVAFLAVITVNRFOAVTRPIKTAQATTKRGFL 136
QY 232 TCGLVNAVTLVMLPFIL-----KQYVLVQPDITTCVHNTCESSPQLYFISLAF 287
DB 137 LSLIIWSIVGAASYFFVLDSTNSPKTKTSGNITRCFEHYE--KGSIPVLIHIFLV-- 192
QY 288 FGLIPFLVLIYCYAAIIRTL-----NA-YDHRWLWYVKASLLILVITFCAPS 337
DB 193 FSEFLVLIILFCNLVLIIRTLTQQVQMQRNAEKRALMWV---CTVLAVFVICFVP-- 247
QY 338 ILIIHH-----ANYNNYNTD---GLYFYLIALCLGSLNSCLDPFLY-FIMSKTRN 385
DB 248 -----HHLVQLPWTALAEGLFQDTHQGINDAHQVTLCLLSTNCVLDPYIYCFLT 302
QY 386 HSTAYLTXXNDLREQQPSQRT 408
DB 303 H-----LTEKLYSMRESRKCSRAT 321

RESULT 11
Q9TTY5 PRELIMINARY; PRT; 342 AA.
AC Q9TTY5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Platelet-activating factor receptor.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;

RN [1]
RP SEQUENCE FROM N.A.
RA Yang W.S., Diehl J.R., Roudabush W.E.;
RT "Partial Sequence of Bovine Platelet-Activating Factor Receptor
RT Gene.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC [2]
CC SEQUENCE FROM N.A.
CC He B., Tiemann U., Kanitz W., Weikard R., Laurent P., Schwerin M.,
CC Schmidt P.;
CC "Molecular characterization of bovine platelet-activating factor
CC receptor transcripts and their detection in different tissues of
CC cattle.";
CC Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
```

Query Match

```
Db 79 DLLYLTSPLFLIHYASGENWIFGDMCKFIRFSFHNLYSSILFLTCFSIFRYCVIHP 138
Qy 218 FTYRGLPKHTYALVTCGLVWATVFLYMLP--FFILKQYIYVQPDITTCCHDVHNTCESS 275
Db 139 MSCFSIHKTRCAVAVACVAVIISLVAVIPMTFLITSTN----RTRSACLDLTSSDELNT 194
Qy 276 PFQLYYFISLAFFGLIPFVLIIYCYAAIIRTLN-----AYDHRWLWYVVASLIL 326
Db 195 --IKWYNILATTFCLPFLVITLCTYTIITHLTHGLQDSCLOKQAR-----RLTILL 247
Qy 327 VIFTCFAPSNIILIIHANYNYNT---DGLFYIYIALCLGSLNCLDPPFYFLMS 381
Db 248 LAFVCFPLPFLIRVIRIESRLLSISCSIEHQIHEAYIVRPLAADNTFGNLLYVVVS 306

RESULT 14
O46685 PRELIMINARY; PRT; 361 AA.
AC O46685;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Orphan G protein-coupled receptor BRGR1B.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Eutheria; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Ferrer-Martinez A., Felipe A., Casado J., Pastor-Anglada M.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; U88367; AAC05612.1; -.
DR EMBL; U88366; AAC05611.1; -.
DR HSP; P02699; I1F88.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR01564; OGRIRCEPTOR.
DR PRINTS; PR01649; PSYCHOSINER.
DR PROSITE; PS00237; G-PROTEIN RECF F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECF F1.2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
SQ SEQUENCE 361 AA; 40905 MW; 4802E77C9A45DEE1 CRC64;

Query Match 16.1%; Score 347; DB 6; Length 361;
Best Local Similarity 27.4%; Pred. No. 6.1e-22;
Matches 86; Conservative 61; Mismatches 127; Indels 40; Gaps 8;

Qy 104 MGYLTS-----SLSTKLPAIYLLVFWGVGPANATLWMLPRTSICTT-VFY 151
Db 1 MGNITADNTSNCDIDHTIHTLAPVYVWVWLVGFPANCLSYGYLQIKARNEGLVYL 60
Qy 152 TNLAIADFLCVLPFKIAYHLNGNNWVFGVLCRATTIVFYGNMYSILLACISINRY 211
Db 61 CNLTVADLFYICSLPFWLQYVLQHDHWSHDDLSCQVCGILLYENIYISVGFLCCISIDRY 120
Qy 212 LAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQYIYVQPDITTCCHDVHNTC 271
Db 121 LAVAHPRFHQRTLKAMGVSAIIVWKELTSTI-----YFLMHEEVVEDADRHC 172
Qy 272 ESSSPFQ-----LYFISLAFGFLIPFVLIIYCYAAIIRTL-----NAYDHRWLWYV 320
Db 173 FEHVPLEPRQGINRYRL-----VGFLPICLLASVYRGILRAVRSHGTSKQKQIQOR 228
Qy 321 ASLLILVFTICFAPSNIILIIHANYNYNTD---GLFYIYIALCLGSLNCLDPPFY 377
Db 229 LVLSTVIFLACFLPYHVLIVR--SLWESSCDFAKGIFNAYHFSLTSTNSVCVADPVLY 286
Qy 378 FLMSKTRNHSTAYL 391
```

```
Db 287 CFVSETTHRLARL 300
RESULT 15
Q9EP3 PRELIMINARY; PRT; 359 AA.
AC Q9EP3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Angiotensin II type 1 receptor.
GN ATL.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=GALL BLADDER;
RA Busmann A., Noack T., Zucht HD., Forssmann W.G., Meyer M.;
RT "Isolation of Angiotensin I from human hemofiltrate as a potent
RT modulator of KATP channels."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AJ301623; CAC21550.1; -.
DR HSP; P34996; 1DDD.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECF F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECF F1.2; 1.
KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
SQ SEQUENCE 359 AA; 41045 MW; D00F4C2C472DE5CC CRC64;

Query Match 16.0%; Score 345; DB 11; Length 359;
Best Local Similarity 27.6%; Pred. No. 9e-22;
Matches 92; Conservative 58; Mismatches 129; Indels 54; Gaps 9;

Qy 89 CPESA-SHLHVKNATMGYLTSSLTSLKLIPIAYLLVFWGVGPANATVLMWLFRTSIC 146
Db 18 CPKDGHSYIFV-----MIPTLSIIFWVGIFGNSLVIVIFYMVKLV 62
Qy 147 TTVFTYTLAIADFLCVLPFKIAYHLNGNNWVFGVLCRATTIVFYGNMYSILLACI 206
Db 63 ASVFLNLALADICFLTLPLWAVYTAEMRYWPNWFGNVMCKIASVSFNLVSLTCL 122
Qy 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQYIYVQPDITTCCHD 266
Db 123 SIDRYLAIVHPMKSRLRRTMLVAKVTCVIWLMAGLASLPAVIRNVFFIENTNITVCAF 182
Qy 267 VHTCESSPOLYFISLAFFGLIPFVLIIYCYAAIIRTL-NAYD-----HRWLWYV 320
Db 183 HYESQNSTLPIGL--GLTKNILGFMFPFLIILSYTLIWKALKKAYEIQKNKPRNDDIFK 240
Qy 321 ASLLILVIFTCFAPSNIIL-----IIHANYNYNTDGLFYIYIALCLGSLNSC 371
Db 241 IIAIVLFFFSVWPHQIFTLVDLVLQGLIHDCKI-----SDIVDTAMPITICIAYPNNC 296
Qy 372 LDPPFY-----FLMSKTRNHST 388
Db 297 LNPFLYGLCKFKKFKYQLKLVIPPKAKSHST 329

RESULT 16
Q9TTY6 PRELIMINARY; PRT; 296 AA.
AC Q9TTY6;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Platelet-activating factor receptor (Fragment).
```

OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang W.S., Diehl J.R., Murphy K.E.;
 RT "Partial Sequence of Canine Platelet-Activating Factor Receptor
 Gene";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF186831; AA01435.1; -.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; UNKNOWN_1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR_F1_2; 1.
 KW Hypothetical protein; Receptor.
 FT NON TER 1 1
 FT NON TER 296 296
 SQ SEQUENCE 296 AA; 33902 MW; 5494C0771C45E101 CRC64;
 Query Match 15.8%; Score 341.5; DB 6; Length 296;
 Best Local Similarity 30.5%; Pred. No. 1.5e-21;
 Matches 91; Conservative 43; Mismatches 127; Indels 37; Gaps 10;
 QY 115 LIPAIYLLVGVGPANAVTLWM---LFFRTSICITTVYTNLAIALDFLCVTLPPKIAIY 171
 DB 9 LPIVYSIIIFVLGVANSYVWVACLYPSKLNKIEIKFVWNLFWADLLFVTLPLWIIY 68
 QY 172 HLGNNWVGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALV 231
 DB 69 YHNOGNWILPSFLCNLAGCFFFIITYCSVAFLAVITYNRFQAVTRPIKTAQTTRKRGFL 128
 QY 232 TGLVWATVFLYMLPFFILKOEYLVQP-----DITTDVNTCESSPFFOLYFISLAF 287
 DB 129 VSLVIWVAIVAAAYFFIILDTNVVPSKSGNITRCFE-HYKSGVPLVIVHVF1--V 184
 QY 288 FGFLPFLVLIIVCYAAIIRTL-----NA-YDRLWLVYKASLLILVIFTICAPSN 337
 DB 195 FGLFVFLVLIIFCNVVIIRLLMQVLOHNAEVRKRALWV---CIVLAVFIVCFPHH 241
 QY 338 II-----LIHHANYNNTDGLFYIYIALCLGSLNSCLDPFLY-FLMSKTRNH 386
 DB 242 LVQLPWTALAEFGQSSSPH---QGINDAHQVTLCLLSTNCVLDPIIYVCLTKKFRKH 295
 RESULT 17
 Q9ERK9 PRELIMINARY; PRT; 328 AA.
 AC Q9ERK9
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DE 01-JUN-2002 (TREMREL. 21, Last annotation update)
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVEV;
 RX MEDLINE=21160052; PubMed=11259526;
 RA Lazarowski E.R., Rochelle L.G., O'Neal W.K., Ribeiro C.M.P.,
 RA Grubb B.R., Zhang V., Harden T.K., Boucher R.C.;
 RT "Cloning and functional characterization of two murine uridine
 nucleotide receptors reveal a potential target for correcting ion
 transport deficiency in cystic fibrosis gallbladder";
 RL J. Pharmacol. Exp. Ther. 297:43-49 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF298899; AAG24619.1; -.

DR EMBL; BC027331; AA27331.1; -.
 DR HSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 KW Hypothetical protein; Receptor.
 SQ SEQUENCE 328 AA; 36721 MW; 00F9DF5ADADF903E CRC64;
 Query Match 15.8%; Score 340; DB 11; Length 328;
 Best Local Similarity 30.7%; Pred. No. 2.2e-21;
 Matches 90; Conservative 47; Mismatches 122; Indels 34; Gaps 9;
 QY 115 LIPAIYLLVGVGPANAVTLWM---LFFRTSICITTVYTNLAIALDFLCVTLPPKIAIY 173
 DB 28 LTFVSVLVVGVPLNICVIAQICASRRITLTSVAVTLNLAIDLMIYASLPLLIINYA 87
 QY 174 NGNNWVGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALV 232
 DB 88 RGDHWPEGLACRFVRFVYANLHGSILFTLCISFQRYLGICHPASWHKRGRRAAWV 147
 QY 233 CGLVWATVFLYMLPFFILKOEYLVQPDIITDCHDVNTCESSPFF-----OLYFISLAF 288
 DB 148 CGVWMLAVTAQCLPTAVFAATG--IQNRITVCYDL-----SPILSTRYLPYGMALTVI 199
 QY 289 GFLLPFLVLIIVCYAAIIRTLNAYD-----HRLWLVYKASLLILVIFTICAPSNIL 340
 DB 200 GFLPFLVLIIVCYAAIIRTLNAYD-----HRLWLVYKASLLILVIFTICAPSNIL 340
 QY 341 IHHANYNNTDGLFYIYIALCLGSLNSCLDPFLY-FLMSKTR 384
 DB 260 TAYLA---VRSTPGVSCPVLTFATAAYKGRPFASVNSVLDPILFYFTQOKFR 309
 RESULT 18
 Q9NOU1 PRELIMINARY; PRT; 359 AA.
 AC Q9NOU1
 DT 01-OCT-2000 (TREMREL. 15, Created)
 DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
 DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
 DE Angiotensin II type 1 receptor.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADRENAL CORTEX;
 RA Bird I.M., Milligan D.S.;
 RT "Isolation of a Full length Ovine Angiotensin II Type-1 Receptor (ATL-
 R) cDNA";
 RL Endocr. Res. 0:0-0 (2000).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF254119; AA66063.1; -.
 DR HSP; P34996; 1DDD.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Receptor; Transmembrane.
 SQ SEQUENCE 359 AA; 41062 MW; C07010EDB8110EB CRC64;
 Query Match 15.8%; Score 340; DB 6; Length 359;
 Best Local Similarity 30.7%; Pred. No. 2.4e-21;
 Matches 89; Conservative 47; Mismatches 120; Indels 34; Gaps 6;
 QY 115 LIPAIYLLVGVGPANAVTLWM---LFFRTSICITTVYTNLAIALDFLCVTLPPKIAIY 173
 DB 30 MIPTLYSIIFVGLFGNSLVIVFYFMYMLKLTVASVFLNLAIDLCLLTLPLWAVYTA 89

Qy 174 NGNNWVGEVLCRAATTVIFGMYCSILLACISINRYLAIVHPTVYRGFLPKHTYALVTC 233
Db 90 MEYRWPFNYLCKIASGSVSNLYASVFLTLCLSDRYLAIVHMKSRRLRMTLVAKVTC 149
Qy 234 GLVWATVFLYMLPFILKQEVYLVQPDITTCDDVHNTCESSPPQLYFFISLAFPGFLIP 293
Db 150 IIIIWLGLASLPTIIIRNVFFIENTNITVCAFHYEQNSLTPVGL--GLTKNLTGLFLP 207
Qy 294 FVLIYCVAAIIRTL-NAYD-----HRLWVWVKAASLLIVFTICFAPSNIILIIHANY 347
Db 208 FLIILSTYLLWKLTKAYELQKNPKRDKDIFKILAILVLEFFSWPHQIFTFM----- 262
Qy 348 YNNNDGLYFYLL-----IALCLGSLNSCLDPFLYFLMSK 382
Db 263 -----DVLQQLGRDCKIEDIVDTAMPITICLAYFNCLNPLFYGLGK 307
RESULT 19
O97724
ID O97724 PRELIMINARY; PRT; 334 AA.
AC O97724;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CCR5 chemokine receptor (Chemokine receptor 5).
GN CCR5.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Kovacs E.M., Baxter G.D., Robinson W.F.;
RT "Feline peripheral blood mononuclear cells express message for both
CX and CC type chemokine receptors";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-73 FROM N.A.
RA Leutenegger C.M., Mislin C., Lutz H.;
RT "Molecular cloning and expression of the 5' region of the feline
chemokine receptor (CCR) 5";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ009815; CAA0838.1; -
DR EMBL; AF059714; AAC69545.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PROSITE; PS00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; UNKNOWN 1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
FT NON TER 1
SQ SEQUENCE 334 AA; 37594 MW; 695B0FE18DB5532E CRC64;
Query Match 15.7%; Score 338; DB 6; Length 334;
Best Local Similarity 28.2%; Pred. No. 3.4e-21;
Matches 84; Conservative 66; Mismatches 118; Indels 30; Gaps 9;
Qy 111 LSTKLIPALYLLVFGVUPANAVTLWMLF--FRTRSTCTTVFYTNLAIDFLFCVTLPFK 168
Db 28 IAAELPLPYLSLVGSLGFGVGNLLVLLINLCKLRGM-TDYLNLNLAIISDLFLFLFPW 86
Qy 169 IAYHLGNWVGEVLCRAATTVIFGMYCSILLACISINRYLAIVHPTVYRGFLPKHTY 228
Db 87 AHYAANG--WFGDMCKTVTGLTHVGFGNGFFIILLTDVRYLAIVHAFALKARTVTF 144
Qy 229 ALVTGCLVWAFVFLYMLPFFILKQEVYLVQPDITTCDDVHNTCESSPP--POLYFISLA 286
Db 145 GAVTSVTAWAAVAVTSLPGCVFNKLQW-----EDSFYSCRPSPFPGNWTHAVTRS 195
Qy 287 FFGFLIFPVLIIYCAIIRTL-----NAYDHRMLWVWKAASLLIVFTICFAPSNIILI 341
Db 196 VLGLVPLLVLMIVCYSAIRLTLRFLCRNEKKHR---AVKLIFVIMIVYFLWAPNNIVLL 252

Qy 342 IHANYYN-----NTDGLYFIYLIACLGSLNSCLDPFLY-FLMSKTRHSTAYLTR 393
Db 253 LSTFPSEFNSVDCOSTSQLDAQMVTELTGHTCCINPIIYAFVGEFRRLHSVFFRK 310
RESULT 20
O8VE54
ID O8VE54 PRELIMINARY; PRT; 298 AA.
AC O8VE54;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to G protein-coupled receptor 23 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019743; AA19743.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; UNKNOWN 1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
FT NON TER 1
SQ SEQUENCE 298 AA; 33708 MW; A09209FBABCE65F2 CRC64;
Query Match 15.6%; Score 337.5; DB 11; Length 298;
Best Local Similarity 31.3%; Pred. No. 3.3e-21;
Matches 79; Conservative 45; Mismatches 99; Indels 27; Gaps 6;
Qy 147 TTVFYTNLAIDFLFCVTLPFKIAYHLGNWVGEVLCRAATTVIFGMYCSILLACI 206
Db 2 TAIFITNLASDLLFVCTLPFKIFYNFN-RHWPFQDTLCKISGTAFUTNIGSMFLTCI 60
Qy 207 SINRYLAIVHPTVYRGFLPKHTYALVTCGLVWATVFLYMLPFFILKQEVYLVQPDITTCDD 266
Db 61 SVDRFLAIVYPPFRSRTIRTRNSAIVCAGWMLLVSGGISASLF-----STTNVNN 111
Qy 267 VHTCSSSPFQLY--YFISLAF-----GFLIPVLIYCVAAIIR-----TLNAYDHR 314
Db 112 ATTTCFEGFSKRVWKYLSKITIFIEVVGFIIPILNVSSVVLRLRKPATLSQIGTN 171
Qy 315 WLWYKASLLILVFTICFAPSNIILIIHANYNNNDGLY-----FYLIACLGSLN 369
Db 172 KKKVLKMITVHMAVFCVFPYNSVLFLYALVRSQAITNCLLERFAKIMYPITICLATLN 231
Qy 370 SCLDPPFLYF 378
Db 232 CCFDPFIY 240
RESULT 21
O8QGQ4
ID O8QGQ4 PRELIMINARY; PRT; 390 AA.
AC O8QGQ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Type five-like somatostatin receptor.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Carassius.
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Peter R.E.


```
Db 3 PMLKNGYNGWATVTTTAPWAS---LGLSAKTCNNVSPESRIVLVVVYSVACTLGVPA 59
QY 131 NAVTLWMLFRT-RSICITVPTYNLAIAADFLFCVTLFPFKIAYHLNGNNWVGEVLCRAIT 189
Db 60 NCLTAWLALLQVLGNGVLAVALLCALCELLYGTGLPLWVIYIRNQRHWTGLGLACKVTA 119
QY 190 VIFGNNYCSILLIACISINRYLAI VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFI 249
Db 120 YIFCNIVSLLFLCCISCDRFVAVVVALESRRRRRTAILISACIFILVGIVHPVFO 179
QY 250 LKOEYLVQPDITTCCHDVHNTCESSPFQLYYFISLAFPGFLIPFVLIYYCAAIIRT-- 307
Db 180 TE-----DKETCFDM---LQMSRIAGYVARET-VGFAIPLSIIAFTNHRFSIK 227
QY 308 ----LNAVDRHMLWVYKASLLIIVIFTCFAPSNIILIIHHANYNNYNTD-----G 354
Db 228 QSMGLSAAQAKVXH--SAIAVVVIFLVCFAVHLVLLVKAASFYYRGDRNAMCGLEER 285
QY 355 LYFIYLIALCIGLSNCLDPLFLPMSKTRNHSFAYLTKXNDLRE 400
Db 286 LYTASVVFCLSTVNGVADPIIYVLA---DHSRQEVSRHKGWKE 328
```

RESULT 27

```
Q9TV48 PRELIMINARY; PRT; 352 AA.
ID Q9TV48 AC Q9TV48;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CC chemokine receptor type 5.
GN CCR5.
OS Cercopithecus aescanius (black-cheeked white-nosed monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=36223;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=52;
RC STRAIN=52;
RA Muller-Trutwin M.C., Corbet S., Hansen J., Georges-Courbot M.C.,
RA Muller-Trutwin M.C., Corbet S., Hansen J., Fomsgaard A.;
RA Diop O., Rigoulet J., Barre-Sinoussi F., Fomsgaard A.;
RT "Mutations in CCR5-coding sequences are not associated with SIV
RT carrier status in African nonhuman primates.";
RL AIDS Res. Hum. Retroviruses 15:931-939(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=52;
RA Mueller-Trutwin M.C., Corbet S., Hansen J., Diop O.,
RA Georges-Courbot M.-C., Barre-Sinoussi F., Fomsgaard A.;
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035216; AAD44009.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40650 MW; 7906256AA3945266 CRC64;
```

```
Query Match 15.4%; Score 331.5; DB 6; Length 352;
Best Local Similarity 26.4%; Pred. No. 1.3e-20;
Matches 84; Conservative 67; Mismatches 120; Indels 47; Gaps 11;
```

```
QY 106 YLTS-----SLSTKLIPAIYLLVFWGVGPANAVTLMLF--FRTRSICTVFTVNL 154
Db 14 YTTSEPCQKINVKQIARLLPLSLVFI FGVGNILVLLINCKRLKSM-TDIYLLNL 72
QY 155 AIADFLCVTLFPFKIAYHLNGNNWVGEVLCRAITVIFYGNMYSICILLACISINRYLAI 214
Db 73 AISDLLFLITVPFWA--HYAAQWDFGNTMCQLTGLTYFIFGFSIGFIILLITIDRYLAI 130
```

```
QY 215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKOEYLVQPDITTC--DVHNTCE 272
Db 131 VHAFVALKARTVTFGVVTSVITVVAVFASLRPII-----FTTSRERLHYTCS 179
QY 273 SSSPFOLYYP-----ISLAFPGFLIPFVLIYYCAAIIRTL-----NAYDRHMLWYKVA 321
Db 180 SHFPYQYQFWKHFHTKIVILGLVPLVWVVICYSIGILKTLRCRNEKKRHR---AVRL 236
QY 322 SLLILVIFTCAPSNIILIIHHANY--NNTDGLYFIYLIALCIGLSNCLDPL 376
Db 237 IFITMIVYFLWAPYINVLNTFOFFGLNCCSSNRDLQAMQVTTGLMTHCCINPII 296
QY 377 Y-FLMSKTRNHSFAYLTK 393
Db 297 YAFVGEKFRNYLLVFFQK 314
```

RESULT 28

```
Q9Z282 PRELIMINARY; PRT; 382 AA.
ID Q9Z282 AC Q9Z282;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE G protein-coupled receptor G2A.
GN G2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98445372; PubMed=9770487;
RA Weng Z., Fluckiger A.C., Nisbitani S., Wahl M.I., Le L.Q., Hunter C.A.,
RA Fernal A.A., Le Beau M.M., Witte O.N.;
RT "A DNA damage and stress inducible G protein-coupled receptor blocks
RT cells in G2/M.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:12334-12339(1998).
DR EMBL; AF083442; AAC67542.1; -
DR MGD; MGI:1890220; G2a.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01563; G2ARECEPTOR.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 382 AA; 42652 MW; D4F8CE0370CCD610 CRC64;
```

```
Query Match 15.4%; Score 331.5; DB 11; Length 382;
Best Local Similarity 27.2%; Pred. No. 1.4e-20;
Matches 85; Conservative 65; Mismatches 115; Indels 47; Gaps 10;
```

```
QY 101 NATMGYLTSSLSKLIPA-----IYLLVFWGVGPANAVTLMLFPRT- 142
Db 11 NTTLG-VTSVLQSTVPSSETCHVSVEESRVLVVYSAVCLGLPANCITAMLTLLQVL 69
QY 143 -RSICTVFTYNLAIAADFLFCVTLFPFKIAYHLNGNNWVGEVLCRAITVIFYGNMYSICIL 201
Db 70 QRNV-LAVLFLCLSLCELLYISTVPLWIIYIQHKWNLGPQACKVTAYIFFCNIYISIL 128
QY 202 LLACISINRYLAI VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFIK-OEYLVQPD 260
Db 129 LLCCISCDRYMAVVVALESRRGHQRHTAVTISACVILLVGLVNVVPMVKESFCFEP- 187
QY 261 ITTCHDVHNTCESSPFQLYYFISLAFPGFLIPFVLIYYCAAIIIRTNAYDH----RWL 316
Db 188 -----LRNWSKIAGVHLYRFT-FGFAIPGLIATFNHQLFRSILKSDLSLAAQKN 236
QY 317 WYKASLLIIVITICFAPSNIILIIHHA--NYNNYNTDG-----LYFIYLIALCIGS 367
Db 237 KVKRSIAVAVTIFLVCFAVHYVHLVLLVKAASFYYQGDMDAVCAFESRLYTVSMVFLCLST 296
```

Qy 368 LNSCLDFLFLVPL 379
Db 297 VNSVADPIIYVL 308

RESULT 29

Q8TBK4 ID Q8TBK4 PRELIMINARY; PRT; 359 AA.
AC Q8TBK4
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Angiotensin receptor 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022447; AAH22447.1; --
KW Receptor.
SQ SEQUENCE 359 AA; 41087 MW; AF90BE7F21E911BA CRC64;

Query Match 15.3%; Score 331; DB 4; Length 359;
Best Local Similarity 29.0%; Pred. No. 1.4e-20;
Matches 84; Conservative 51; Mismatches 121; Indels 34; Gaps 6;
Qy 115 LIPAIYLLVFGVGPANAVTLMLLFFRTR-SICTVFYTNLAIDFLCVTLPLPKIAYHL 173
Db 30 MIPTLYSIIFVGVFGNSLVIVYFYMKLTVASVFLNLALADLCFLTLPLWAVYA 89
Qy 174 NGNNVGEVLCTATTIVFYGNMVCISILLACISINRYLAIVHPTVGRGLPKHYALVTC 233
Db 90 MEYRPFNGYLCKIASVSNSVFLTCLSIDRYLAIVHPMKSLRRLTMLVAVKTC 149
Qy 234 GLVWATVFLVLPFLKQEVLYVQPDITTDHVDNTCESSPPOLYVFISLAFVGLIP 293
Db 150 IIIWLLAGLASLPALIHNVFFIENTITVCAFHYESQNSTLPIGL--GLTKNIGLFLP 207
Qy 294 FVLIIYCYAAIIRTL-NAYD-----HRLWVYKASLLILVITFCFAPSNIILIHANY 347
Db 208 FLIILTSYTLIAKAVKAYEQKNPRNDIFKIIMAILVLPFPSPWIPHOIFTEL----- 262
Qy 348 YNNTDGLYFIYL-----IALCLGSLNSCLDPPFLYFLMSK 382
Db 263 -----DVLIIQLGIIRDCRIADIVDTAMPITICIAFYFNCLNPLFYGLGK 307

RESULT 30

Q9TUR6 ID Q9TUR6 PRELIMINARY; PRT; 339 AA.
AC Q9TUR6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (fragment).
GN CCR5.
OS Cercopithecus aethiops vervet.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=100936;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wollinkay S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.

EMBL; AF162026; AAD47781.1; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF0001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 339 339
SQ SEQUENCE 339 AA; 39178 MW; 9DF2A6FA446C55AED CRC64;
Query Match 15.3%; Score 330.5; DB 6; Length 339;
Best Local Similarity 25.8%; Pred. No. 1.5e-20;
Matches 82; Conservative 69; Mismatches 124; Indels 43; Gaps 10;
Qy 104 MGYLTS-----SLSTKLIPAIYLLVFGVGPANAVTLMLF--FRTRSICTTVFYT 152
Db 5 INYITSEPCOKINVKQIAARLLPLLSLVIFIFGVGNILVLLINCKRLKSM-TDIYLL 63
Qy 153 NLAIADFLFCVTLPLPKIAYHLNGNNVFGVGLVCRATTIVFYGNMVCISILLACISINRYL 212
Db 64 NLAISDLLFLTLVFPWA--HYAAAQWDFGNTMCQLLTGLYFIFGFGIFILLITIDRYL 121
Qy 213 AIVHPPTVGRGLPKHYALVTCGLVWATVFLYMLPFFILKQEVLYVQPDITTDHVDNTCE 272
Db 122 AIVHAVFALKARTVTTEGVTSVITWVAVFASLPRIIFTR-----SOREGLHYTCS 172
Qy 273 SSSPFFOLYF-----ISLAFVGLIPFLVLIYCYAAIIRTL-----NAYDHRWLWYVKA 321
Db 173 SHFPYQYQWKNFQILKIVILGLVLPVLMVWICYSGILKTLRCNEKKRHR---AVRL 229
Qy 322 SLILVITFCFAPSNIILIHANY-----NNTDGLYFIYLIACLSLNSCLDPPFL 376
Db 230 IFTIMIVYFLWAPYINVLNLTTFQEFFGLNCCSSNRLDQAMQVTELTGWTCCIDPII 289
Qy 377 Y-FLMSKTRHSTAYLTK 393
Db 290 YAFVGEKFRNYLLVFFQK 307
RESULT 31
Q93239 ID Q93239 PRELIMINARY; PRT; 342 AA.
AC Q93239;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CXC chemokine receptor-2.
OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99367403; PubMed=10436187;
RA Fujiki K., Shin D., Nakao M., Yano T.;
RT "Molecular cloning of carp (Cyprinus carpio) CC chemokine, CXC
RT chemokine receptors, allograft inflammatory factor-1, and natural
RT killer cell enhancing factor by use of suppression subtractive
RT hybridization";
RL Immunogenetics 49:909-914(1999).
DR EMBL; AB010713; BAA31470.1; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR01559; DUFFYANTIGEN.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
SQ SEQUENCE 342 AA; 38481 MW; 8946E5ED5E534B39 CRC64;
Query Match 15.3%; Score 330; DB 13; Length 342;

```
Best Local Similarity 27.8%; Pred. No. 1.7e-20;
Matches 82; Conservative 57; Mismatches 122; Indels 34; Gaps 9;

Qy 119 IYLVFVGVGPANAVTLM-L-FPRTSICITVFTVNTLAIADFLFCVTLPKIAYHLNGNN 177
Db 32 VYVSFCLSPGNWVIFVWSCMENRRSTDDVLMHLAIADFLFALTLPFSAA-DIHAGH 90

Qy 178 WVGVEVLCRATTVIFYGNMYCSILLACISINRYLAIVHFTYRGLPKHYALVTCGLVW 237
Db 91 WMFTGFWCKLISGQEAFTFCCVPELLACISVDRLAIVKATQFLAQRHLVGLV-CALVW 149

Qy 238 ATVELYMLPFFILKQEYLVQPDITTCCHDVHNTCESSPFQLYFYISLAFPGFLIPFVLI 297
Db 150 VCAFLLSLPVNVQAFITGNMGDYICH-DNVTAESMDSWRMGLRIHLHTLFGFLPLAVM 208

Qy 298 IYCY-----AAIIRTLNAYDRHRLWYKASLLILVIFTCFAPSNIILIIHHANYNNT- 352
Db 209 MFCYGFTMCTLCRTNSQKQK---AMRVILSVVLAFTVCLWLPFNILELI-----DTL 257

Qy 353 -----DGLYFIYLIACLSLNSCLDPFLY-FLMSKTRNHSTAYLTk 393
Db 258 MRGKVIETCELDRSIDVALYVTOAMAFHCAINPILYAFIGKFRNQLLMSLFK 312

RESULT 32
Q9TUUS PRELIMINARY; PRT; 339 AA.
ID Q9TUUS
AC Q9TUUS;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.
OS Macaca fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9542;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprende J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162032; AAD47787.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
FT NON TER 1
FT NON TER 339
SQ SEQUENCE 339 AA; 39124 MW; EA9CD7AA531E599A CRC64;

Query Match 15.3%; Score 329.5; DB 6; Length 339;
Best Local Similarity 25.9%; Pred. No. 1.8e-20;
Matches 82; Conservative 68; Mismatches 123; Indels 43; Gaps 10;

Qy 106 YLTS-----SLSTKLIPALYLLVFGVGPANAVTLMWLF--FRTRSICTTVFTVNTL 154
Db 7 YTTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNILVLLINCKRLKSM-TDIYLLNL 65

Qy 155 AIADFLFCVTLPPFKIAYHLNGNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAI 214
Db 66 AISDLLFLTLVFPWA--HYAAAQWDFGNTMQLTGLYFIFGFSGIFILLITDRYLAI 123

Qy 215 VHPFTYRGLPKHYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITTCCHDVHNTCESS 274
Db 124 VHAFVALKARTVFGVTSVITWVAVFASLPGLIFTR-----SOREGLHYTCSSH 174

Qy 275 SPFQLYYF-----ISLAFPGFLIPFVLIICYAAIIRTL-----NAYDRHRLWYKASL 323
Db 175 FFSQVQFQWKNFRTKIVILGLVLPVLMVVICSGILKTLRLCRNEKKRHR---AVRLIF 231

Qy 324 LILVITFCFAPSNIILIIHHANY--NNTDGLYFIYLIACLSLNSCLDPFLY- 377
Db 232 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRDLQAMQVTTGLMTHCCINPIYA 291

Qy 378 FLMSKTRNHSTAYLTk 393
Db 292 FVGEKFRNLYLVFFQK 307
```

```
175 FFSQVQFQWKNFRTKIVILGLVLPVLMVVICSGILKTLRLCRNEKKRHR---AVRLIF 231
324 LILVITFCFAPSNIILIIHHANY--NNTDGLYFIYLIACLSLNSCLDPFLY- 377
232 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRDLQAMQVTTGLMTHCCINPIYA 291
378 FLMSKTRNHSTAYLTk 393
292 FVGEKFRNLYLVFFQK 307

RESULT 33
Q9TUR4 PRELIMINARY; PRT; 339 AA.
ID Q9TUR4
AC Q9TUR4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.
OS Mandrillus sphinx (Mandrill) (Papio sphinx).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Mandrillus.
OX NCBI_TaxID=9561;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprende J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162032; AAD47787.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
FT NON TER 1
FT NON TER 339
SQ SEQUENCE 339 AA; 518626A2DEF1BFBE CRC64;

Query Match 15.3%; Score 329.5; DB 6; Length 339;
Best Local Similarity 25.9%; Pred. No. 1.8e-20;
Matches 82; Conservative 69; Mismatches 122; Indels 43; Gaps 10;

Qy 106 YLTS-----STKLIPALYLLVFGVGPANAVTLMWLF--FRTRSICTTVFTVNTL 154
Db 7 YTTSEPCQKINVKQIAARLLPPLYSLVFIFGVGNILVLLINCKRLKSM-TDIYLLNL 65

Qy 155 AIADFLFCVTLPPFKIAYHLNGNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAI 214
Db 66 AISDLLFLTLVFPWA--HYAAAQWDFGNTMQLTGLYFIFGFSGIFILLITDRYLAI 123

Qy 215 VHPFTYRGLPKHYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITTCCHDVHNTCESS 274
Db 124 VHAFVALKARTVFGVTSVITWVAVFASLPGLIFTR-----SOREGLHYTCSSH 174

Qy 275 SPFQLYYF-----ISLAFPGFLIPFVLIICYAAIIRTL-----NAYDRHRLWYKASL 323
Db 175 FFSQVQFQWKNFRTKIVILGLVLPVLMVVICSGILKTLRLCRNEKKRHR---AVRLIF 231

Qy 324 LILVITFCFAPSNIILIIHHANY--NNTDGLYFIYLIACLSLNSCLDPFLY- 377
Db 232 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRDLQAMQVTTGLMTHCCINPIYA 291

Qy 378 FLMSKTRNHSTAYLTk 393
Db 292 FVGEKFRNLYLVFFQK 307
```

```
RESULT 34
Q95NC3 PRELIMINARY; PRT; 352 AA.
AC Q95NC3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE C-C chemokine receptor 5.
GN CCR5.
OS Micropithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Y., Ryder O.A., Zhang Y.;
RT "Sequence comparison of the CCR5 gene in primates and primate
RT phylogeny.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF177886; AAK43369.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 352 AA; 40546 MW; 6464152F3E566AE5 CRC64;

Query Match 15.3%; Score 329.5; DB 6; Length 352;
Best Local Similarity 25.8%; Pred. No. 1.9e-20;
Matches 82; Conservative 69; Mismatches 124; Indels 43; Gaps 10;

QY 104 MGYLT-----SLSTKLIPAIYLLVFGVGPANAVTLMLF--FRTRSICTVFT 152
D 12 INYTSPECQKINVKQIAARLLPLYSLVFIFGVGNILVLLINCKRLKSM-TDIYLL 70
QY 153 NLAIADFLFCVTLFPKIAIYHLNGNNVFGVLCRATTVIFYGNMYCSILLACISINRYL 212
D 71 NLAISDLLFLTLVPEWA--HYAAQAQWDFGNTMCRLLTGLYFIFGFGIFILLTIDRYL 128
QY 213 AIHVPFYRGPKHYALVTCGLWATVFLYMLPFILKQEYLVQPDITTCCHDVNTCE 272
D 129 AIHVAFAKARTVTFGVVTSVITWVAVFASLPGLIIFTR-----SQREGLHYTCS 179
QY 273 SSSPFQLYF-----ISLAFEGFLIPFVLIICYAAIIRTL-----NAVDRHLMVYVKA 321
D 180 SHFPYQYQFKNFQTLKIVILGLVPLVWVICYSGILKTLRCRNEKKRHR---AVRL 236
QY 322 SLLILVFTICFAPSNIILIIHHANY--NNTDGLYFYIYIALCLGSLNSCLDPFL 376
D 237 IFTIMIVFLWAPYNIIVLLNTFQEFFGLNCSNRLDQAMQVETLGLMTHCCINPII 296
QY 377 Y-FLMSKTRNHSTAYLTK 393
D 297 YAFVGEKFRNLYLVFFQK 314

RESULT 35
Q9TUU8 PRELIMINARY; PRT; 339 AA.
AC Q9TUU8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.
OS Cercopithecus diana (Diana monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=36224;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF161949; AAD47705.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 339
SQ SEQUENCE 339 AA; 39049 MW; 6D1A93F66270F3ED CRC64;

Query Match 15.2%; Score 328.5; DB 6; Length 339;
Best Local Similarity 25.9%; Pred. No. 2.2e-20;
Matches 82; Conservative 69; Mismatches 122; Indels 43; Gaps 10;

QY 106 YLTS-----SLSTKLIPAIYLLVFGVGPANAVTLMLF--FRTRSICTVFTYTNL 154
D 7 YTTSEPCQKINVKQIAARLLPLYSLVFIFGVGNILVLLINCKRLKSM-TDIYLLNL 65
QY 155 AIADFLFCVTLFPKIAIYHLNGNNVFGVLCRATTVIFYGNMYCSILLACISINRYLAI 214
D 66 AISDLLFLTLVPEWA--HYAAQAQWDFGNTMCRLLTGLYFIFGFGIFILLTIDRYLAI 123
QY 215 VHPFTYRGPKHYALVTCGLWATVFLYMLPFILKQEYLVQPDITTCCHDVNTCESS 274
D 124 VHAFAKARTVTFGVVTSVITWVAVFASLPGLIIFTR-----SQREGLHYTCSH 174
QY 275 SPFQLYF-----ISLAFEGFLIPFVLIICYAAIIRTL-----NAVDRHLMVYVKA 323.
D 175 FPFYQYQFKNFQTLKIVILGLVPLVWVICYSGILKTLRCRNEKKRHR---AVRLIF 231
QY 324 LILVFTICFAPSNIILIIHHANY--NNTDGLYFYIYIALCLGSLNSCLDPFLY- 377
D 232 TIMIVFLWAPYNIIVLLNTFQEFFGLNCSNRLDQAMQVETLGLMTHCCINPIYA 291
QY 378 FLMSKTRNHSTAYLTK 393
D 292 FVGEKFRNLYLVFFQK 307

RESULT 36
Q9TSQ3 PRELIMINARY; PRT; 339 AA.
AC Q9TSQ3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstman K., Chen Z., Korber B., Oprendeck J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162007; AAD47762.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PSS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
```

Db	31	SKYFLPCMYLVFVCGVGNLSLVISIFYHKQSLSTDVFLVNLPLADLVFVCTLPFFWA- 89
QY	171	YHLGNNGWVGEVLCRATTWIFYGNMYCSILLACISINRYLAIVHPF--TVRGLPGH-TY 228
Db	90	-YAGIHEWIEGQVMCKTLGLGYTTFNFTYSMLILTCITVDREIFVVVKATKAYNQQAQRMTW 148
QY	229	ALVTCGLVMAVTVELMYLPFFFLKQEYLYVQPDITTC--HDVHNTCESSPFOLYYFISLA 286
Db	149	GKVICLLIWISSLVSLPQIIYGNVFNL---DKLICRYHD-----EELSTVVLATQMTL- 199
QY	287	FFGFLIPFVLIIYCYAAIIRTL-----NAYDRHWLYWKASLLILVIFPICFAPSNIILII 342
Db	200	--GFFLPLLTMIVCYSVIIITKLLHAGGFQKHSRSL---KIIFLVMVAVFLTQTFFPNLVKLI 254
QY	343	HHANYNNYNDGLFYVILIALCLGSLNSCLDPFFLYFLMS 381
Db	255	RSTHWEYYAMTSPHTYIIIVTEAIYALRACLNPFVLYAFVS 293

RESULT 38	
Q9MYJ8	
ID	PRELIMINARY; PRT; 355 AA.
AC	Q9MYJ8;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE	Chemokine receptor.
GN	CCR1.
OS	Callithrix jacchus (Common marmoset).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callithrix.
OX	NCBI_TaxID=9483;
PN	[1]

```

[1]
SEQUENCE FROM N.A.
MEDLINE=20153429; PubMed=10686294;
RA Liang M., Rosser M.P., Ng H.P., May K., Bauman J.G., Islam I.,
RA Ghanam A., Kreschmer P.J., Pu H., Dunning L., Snider R.M.,
RA Morrissey M.M., Hesselgeser J., Perez H.D., Horuk R.;
RT "Species selectivity of a small molecule antagonist for the CCR1
RT chemokine receptor.";
RL Eur. J. Pharmacol. 389:41-49(2000).
DR EMBL: AF127528; AAF36453.1; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7cm_1; 1.
DR PRINTS: PR00237; GPCRRHODOPSN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
DR PROSITE: PS0262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 355 AA; 40928 MW; 2B01C47E9874A2C1 CRC64;
Query Match 15.2%; Score 328; DB 6; Length 355;
Best Local Similarity 28.1%; Pred. No. 2.6e-20;
Matches 82; Conservative 63; Mismatches 121; Indels 26; Gaps 10
Qy 110 SLSTKLIPAIYLLVFGVPANAVTLMLF-FRTRSICTTVFYTNLAIDFLCVTLDPK 168
Db :|::||::||::||::||::||::||::||::||::||::||::||::||::||
31 AFGAKLLPPLYSLVFVLGVNLVVVVVQVKLKNNMTSYILLNLSDLLFLFTLPW 90
Qy 169 IAYHLGNWVFGEVLCRAITVIFGNNMYCSILLIACISINRYLAIVHPFTYRGLPKH 228
Db :|::||::||::||::||::||::||::||::||::||::||::||::||
91 ISYQLK-TDMWFVGAMCKVLSGFYTGYSEIFFILITDIRYLAIVHAVFAALARTVT 149
Qy 229 ALVTCGLVAWATVELYMLP-PFILKOEYLVQPDIITCHDVNTC-----ESSSPFOLY 281
Db :|::||::||::||::||::||::||::||::||::||::||::||::||
150 GVITTSIIIVLUAILASLGILFAKTQW-----EIT-----HRTCSLHFPHESRQEWLFQ 199
Qy 282 FISLAFFGLPIFFVIIICYAAIIR-TLNAYDHRWLWVYKASLLILVIETICFAPSN--- 337
Db :|::||::||::||::||::||::||::||::||::||::||::||::||
200 AKLNILLGLVPLLVMIVCYGTGIKILLRRPNKKSKAVRLFIVIMIIFLFWTPYNLT 259
Qy 338 ILIHIANHYNNYNDTG--LVFIYLIALCLGSNSCLDDPFY-FLMSKTRNH 386
Db :|::||::||::||::||::||::||::||::||::||::||::||::||

```



```
Db 260 LISVFQDFLFTYGCQGRLDLAIQVTEMIAYTHCCVNPVIYAFVGERFRKH 311
RESULT 39
Q9TUU6 PRELIMINARY; PRT; 339 AA.
AC Q9TUU6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.
OS Macaca fuscata (Japanese macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9542;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprende J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF161956; AAD47712.1; -.
DR EMBL; AF161956; AAD47753.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 1
FT NON TER 339
SQ SEQUENCE 339 AA; 39080 MW; 4DCA9C72B093216D CRC64;

Query Match 15.2%; Score 327.5; DB 6; Length 339;
Best Local Similarity 25.9%; Pred. No. 2.7e-20;
Matches 82; Conservative 68; Mismatches 123; Indels 43; Gaps 10;

Qy 106 YLTS-----SLSTKLIPALYLLVFGVGPANAVTLMLF--FRTRSICTTVFVTLN 154
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
7 YTTSEPCQKINVKQIAARLLPPLSLVFIFFGVGNILVLLINCKRLKSM-TDIYLLNL 65

Qy 155 AIADFLFCVTLFPFKIAYHLNGNNVFGVLCRATTVIFYGNMYCSILLACISINRYLAI 214
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
66 AISDLLFLTLVFWA--HYAAARWDFGNTMCQLLTGLYFIFGFSGIFPIILLTIDRYLAI 123

Qy 215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITTHDVHNTCESS 274
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
124 VHAFVALKARTVTFGVVTSVITWVAVFASLPGLIIFTR-----SQRGLHYTCSSH 174

Qy 275 SPQLYYF-----ISLAFFGLIPFLVLIICYAAIIRTL-----NAVDRHLMVVKASL 323
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
175 FPYSQYQFQKFNFTLKIVILGLVPLLVWVICYSGILKTLRCRNEKKRHR---AVRLIF 231

Qy 324 LILVIFTICFAPSNIILIIHANYYY-----NNTDGLYFIYLIACLGSLNSCLDPFLY- 377
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
232 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRDLQAMQVTTGLMTHCCINPIYA 291

Qy 378 FLMSKTRNHSTAYLTK 393
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
292 FVGEKFRNLYLVFFQK 307

RESULT 40
Q9TQX3 PRELIMINARY; PRT; 339 AA.
AC Q9TQX3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE C-C chemokine receptor 5 (Fragment).
GN CCR5.

OS Mandrillus sphinx (Mandrill) (Papio sphinx).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Mandrillus.
OX NCBI_TaxID=9561;
RN [1]
RP SEQUENCE FROM N.A.
RA Kunstan K., Chen Z., Korber B., Oprende J., Stanton J., Agy M.,
RA Shibata R., Yoder A., Pillai S., Kuiken C., Marx P., Wolinsky S.;
RT "Sequences of the CCR5 genes from diverse simian and prosimian
RT species.";
RT Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF161999; AAD47754.1; -.
DR EMBL; AF161998; AAD47753.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 1
FT NON TER 339
SQ SEQUENCE 339 AA; 39119 MW; 066BFC06CD00F0D8 CRC64;

Query Match 15.2%; Score 327.5; DB 6; Length 339;
Best Local Similarity 25.6%; Pred. No. 2.7e-20;
Matches 81; Conservative 70; Mismatches 122; Indels 43; Gaps 10;

Qy 106 YLTS-----SLSTKLIPALYLLVFGVGPANAVTLMLF--FRTRSICTTVFVTLN 154
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
7 YTTSEPCQKINVKQIAARLLPPLSLVFIFFGVGNILVLLINCKRLKSM-TDIYLLNL 65

Qy 155 AIADFLFCVTLFPFKIAYHLNGNNVFGVLCRATTVIFYGNMYCSILLACISINRYLAI 214
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
66 AISDLLFLTLVFWA--HYAAARWDFGNTMCQLLTGLYFIFGFSGIFPIILLTIDRYLAI 123

Qy 215 VHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITTHDVHNTCESS 274
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
124 VHAFVALKARTVTFGVVTSVITWVAVFASLPGLIIFTR-----SQRGLHYTCSSH 174

Qy 275 SPQLYYF-----ISLAFFGLIPFLVLIICYAAIIRTL-----NAVDRHLMVVKASL 323
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
175 FPYSQYQFQKFNFTLKIVILGLVPLLVWVICYSGILKTLRCRNEKKRHR---AVRLIF 231

Qy 324 LILVIFTICFAPSNIILIIHANYYY-----NNTDGLYFIYLIACLGSLNSCLDPFLY- 377
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
232 TIMIVYFLWAPYNIIVLLNTFQEFFGLNCCSSNRDLQAMQVTTGLMTHCCINPIYA 291

Qy 378 FLMSKTRNHSTAYLTK 393
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
292 FVGEKFRNLYLVFFQK 307

Search completed: June 24, 2003, 12:05:38
Job time : 54.5632 secs
```

HIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:53:16 ; Search time 14.5178 Seconds
(without alignments)
1165.628 Million cell updates/sec

Title: US-09-208-629F-6

Perfect score: 2157

Sequence: 1 CSMILQISXRLRDGTQVIM.....AYLTXXNDLREQQPSQRT 408

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1998	92.6	374	1 PAR3 HUMAN	O00254 homo sapien
2	1455.5	66.1	369	1 PAR3 MOUSE	O08675 mus musculus
3	1414	65.6	368	1 PAR3 RAT	Q020e1 rattus norv
4	606.5	28.1	399	1 PAR2 MOUSE	P55086 mus musculus
5	599.5	27.8	420	1 PAR1 XENLA	P47749 xenopus lae
6	595	27.6	397	1 PAR2 RAT	O63645 rattus norv
7	589.5	27.3	397	1 PAR2 HUMAN	P55085 homo sapien
8	535	24.8	395	1 PAR4 RAT	Q020e0 rattus norv
9	531.5	24.6	396	1 PAR4 MOUSE	O08634 mus musculus
10	522.5	24.2	430	1 PAR1 MOUSE	P30558 mus musculus
11	521.5	24.2	385	1 PAR4 HUMAN	Q96r10 homo sapien
12	520.5	24.1	425	1 PAR1 PAPH	P56488 papio hamad
13	519	24.1	425	1 PAR1 HUMAN	P55116 homo sapien
14	505.5	23.4	432	1 PAR1 RAT	P26824 rattus norv
15	501	23.2	428	1 PAR1 CRIL	Q00991 cricetus
16	428	19.8	537	1 P2Y8 XENLA	P79928 xenopus lae
17	426	19.7	362	1 P2Y8 MELGA	P49652 meleagris g
18	424	19.7	362	1 P2YR CHICK	P34996 gallus gall
19	410.5	19.0	361	1 EB12 HUMAN	P32249 homo sapien
20	410.5	19.0	367	1 GP17 HUMAN	Q13304 homo sapien
21	409.5	19.0	373	1 P2YR RAT	P49651 rattus norv
22	406.5	18.8	308	1 P2Y5 CHICK	P32250 gallus gall
23	399	18.5	373	1 P2YR HUMAN	P47900 homo sapien
24	398	18.5	373	1 P2YR BOVIN	P48042 bos taurus
25	397.5	18.4	373	1 P2YR MOUSE	P49650 mus musculus
26	384.5	17.8	370	1 P2Y9 HUMAN	Q99677 homo sapien
27	382	17.7	365	1 P2Y4 HUMAN	P51582 homo sapien
28	380.5	17.6	309	1 CLT2 RAT	Q024t9 rattus norv
29	379.5	17.6	345	1 CLT2 PIG	Q95n03 sus scrofa
30	378	17.5	361	1 P2Y4 RAT	Q05811 rattus norv
31	376.5	17.5	346	1 CLT2 HUMAN	Q9ns75 homo sapien
32	376	17.4	344	1 P2Y5 HUMAN	P43657 homo sapien
33	375.5	17.4	340	1 CLT1 PIG	Q95n02 sus scrofa

RESULT 1

PAR3_HUMAN
ID PAR3_HUMAN STANDARD; PRT; 374 AA.
AC O00254;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-like 2) (Coagulation factor II receptor-like 2).
GN F2RL2 OR PAR3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RX SEQUENCE FROM N.A., AND MUTAGENESIS OF THR-39 AND PHE-40.
RA Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W., Timmons C., Tram T., Coughlin S.R.;
RT "Protease-activated receptor 3 is a second thrombin receptor in humans.";
RL Nature 386:502-506(1997).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANTS SER-15; VAL-177 AND ASP-250.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q., Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP TISSUE SPECIFICITY.
RX MEDLINE=98279023; PubMed=9614115;
RA Schmidt V.A., Nierman W.C., Maglott D.R., Cupit L.D., Moskowitz K.A., Wainer J.A., Bahou W.F.;
RT "The human proteinase-activated receptor-3 (PAR-3) gene. Identification within a PAR gene cluster and characterization in vascular endothelial cells and platelets.";
RL J. Biol. Chem. 273:15061-15068(1998).
RN [4]
RP FUNCTION.
RX MEDLINE=99178892; PubMed=10079109;
RA Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H., Coughlin S.R.;
RT "Protease-activated receptors 1 and 4 mediate activation of human platelets by thrombin.";
RL J. Clin. Invest. 103:879-887(1999).
CC -!- FUNCTION: Receptor for activated thrombin coupled to G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highest expression in the megakaryocytes of the bone marrow, lower in mature megakaryocytes, in platelets and in a variety of other tissues such as heart and gut.
CC -!- PWM: A proteolytic cleavage generates a new amino terminus that functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

ALIGNMENTS

34	372.5	17.3	337	1	CLT1_HUMAN
35	371.5	17.2	352	1	CLT1_MOUSE
36	370	17.2	361	1	P2Y4_MOUSE
37	366	17.0	377	1	P2Y2_HUMAN
38	365.5	16.9	339	1	CLT1_RAT
39	365.5	16.9	374	1	P2Y2_RAT
40	365.5	16.9	381	1	GP34_HUMAN
41	364.5	16.9	328	1	P2Y3_CHICK
42	363.5	16.9	355	1	IL8A_RABIT
43	361.5	16.8	309	1	CLT2_MOUSE
44	361.5	16.8	363	1	AG2S_XENLA
45	361	16.7	373	1	P2Y2_MOUSE

Q9y271 homo sapien
Q99ja4 mus musculus
Q9jje7 mus musculus
P41231 homo sapien
Q924t8 rattus norv
P41232 rattus norv
Q98905 homo sapien
Q98907 gallus gall
P21109 oryctolagus
Q920a1 mus musculus
P35373 xenopus lae
P35383 mus musculus

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U92971; AAC51218.1; --
 EMBL; AF374726; AAK51564.1; --
 HSP; P34996; 1DDO
 Genew; HGNC:3539; F2RL2.
 MIN; 601919; --
 InterPro; IPR000276; GPCR_Rhodpsn.
 Pfam; PF00001; 7tm_1; 1.
 PRINTS; PR00237; GPCRHOOPS.
 PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 PROSITE; PS00262; G_PROTEIN_RECP_F2_1; 1.
 G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Blood coagulation; Polymorphism.
 SIGNAL 1 21 POTENTIAL.
 PROPEP 22 38 REMOVED FOR RECEPTOR ACTIVATION (BY
 SIMILARITY)
 CHAIN 39 374 PROTEINASE ACTIVATED RECEPTOR 3..
 DOMAIN 39 94 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 95 120 1 (POTENTIAL).
 DOMAIN 121 128 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 129 148 2 (POTENTIAL).
 DOMAIN 149 167 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 168 189 3 (POTENTIAL).
 DOMAIN 190 206 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 207 230 4 (POTENTIAL).
 DOMAIN 231 259 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 261 280 5 (POTENTIAL).
 DOMAIN 281 297 CYTOPLASMIC (POTENTIAL).
 TRANSMEM 298 322 6 (POTENTIAL).
 DOMAIN 323 336 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 337 361 7 (POTENTIAL).
 DOMAIN 362 374 CYTOPLASMIC (POTENTIAL).
 SITE 38 39 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
 DISULFID 166 245 BY SIMILARITY.
 CARBOHYD 25 25 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 331 331 N-LINKED (GLCNAC. .) (POTENTIAL).
 VARIANT 15 15 L -> S.
 M -> V. /FTId=VAR_012849.
 M -> D. /FTId=VAR_012850.
 T -> P: NO PROTEOLYTIC CLEAVAGE (BY THROMBIN).
 E -> A: ALTERED SIGNAL UPON THROMBIN CLEAVAGE.
 Query Match 92.6%; Score 1998; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 4.4e-102;
 Matches 374; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 20 MKALIFAAAGLLLLPTFCQSGMENDTNNAKPTLPKTPRGAPNPFEEPPFSALEGWT 79
 1 MKALIFAAAGLLLLPTFCQSGMENDTNNAKPTLPKTPRGAPNPFEEPPFSALEGWT 60
 80 GATITVKIKCEESASHLVKVNATNGYLTSSLSKLPKIPAILLVVGVGVPANATLWMLF 139
 61 GATITVKIKCEESASHLVKVNATNGYLTSSLSKLPKIPAILLVVGVGVPANATLWMLF 120
 140 FRTSICITVFTNLAIDFLPCVTLTPFKIAYHLNGNNWVGEVLCRATTIVIFGNNMYCS 199
 121 FRTSICITVFTNLAIDFLPCVTLTPFKIAYHLNGNNWVGEVLCRATTIVIFGNNMYCS 180
 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQP 259
 181 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQP 240

QY 260 DITTCDDVHNTCESSPPFQLYYFISLAFFGLIPPLVLIICVAAIIRTNAYDHRWLWYV 319
 DB 241 DITTCDDVHNTCESSPPFQLYYFISLAFFGLIPPLVLIICVAAIIRTNAYDHRWLWYV 300
 QY 320 KASLLILVFTICFAPSNIILIIHHANYNNYNTDGLYFYIYLIATLCGLSGLNSCLDPFLYFL 379
 DB 301 KASLLILVFTICFAPSNIILIIHHANYNNYNTDGLYFYIYLIATLCGLSGLNSCLDPFLYFL 360
 QY 380 MSKTRNHSTAYLTK 393
 DB 361 MSKTRNHSTAYLTK 374
 RESULT 2
 PAR3_MOUSE
 ID PAR3_MOUSE STANDARD; PRT; 369 AA.
 AC O08675;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-
 DE like 2) (Coagulation factor II receptor-like 2).
 GN F2RL2 OR PAR3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6;
 RX MEDLINE=97242411; PubMed=9087410;
 RA Ishihara H., Connolly A.J., Zeng D., Kahn M.L., Zheng Y.-W.,
 RA Timmons C., Tram T., Coughlin S.R.;
 RT "Protease-activated receptor 3 is a second thrombin receptor in
 RT humans.";
 RL Nature 386:502-506(1997).
 CC -1- FUNCTION: High affinity receptor for activated thrombin coupled to
 CC G proteins that stimulate phosphoinositide hydrolysis. May play a
 CC role in platelets activation.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
 CC functions as a tethered ligand.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
FT DOMAIN 189 205 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 206 229 4 (POTENTIAL).
FT DOMAIN 230 258 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 260 279 5 (POTENTIAL).
FT DOMAIN 280 296 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 297 321 6 (POTENTIAL).
FT DOMAIN 322 335 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 336 369 7 (POTENTIAL).
FT DOMAIN 361 369 CYTOPLASMIC (POTENTIAL).
FT SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 165 244 BY SIMILARITY.
FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 369 AA; 41707 MW; F416GBF766D073DB CRC64;

Query Match 66.1%; Score 1425.5; DB 1; Length 369;
Best Local Similarity 71.6%; Pred. No. 4.4e-71;
Matches 260; Conservative 44; Mismatches 58; Indels 1; Gaps 1;

Qy 20 MKALIFAAAGLLLLPTFCQSGMENDTNNAKPTLPKTRFGAPNSFEFPFSALEGWT 79
Db 1 MKLILVAAGLLFLPVTVCOSGI-NVSDNSAKPTLTIKSFNGGPONTFEBPPLSDIEGWT 59

Qy 80 GATITVKIKPEESASHLVKNATMGVLTSSLSKLPALYLLVGVGVPANAVTLWLF 139
Db 60 GATTIKAECPEDSISTLVNNAATGYLRSLSQVPAIYILLVGVGVPNSIVTLWKL 119

Qy 140 FRTSICITVFTYTNLADELFCVTLPFKIAYHLNGNNWVGEVLCRATTIVFGNNMYS 199
Db 120 LRTKSIISLVFHTNLATDLFCVTLPFKIAYHLNGNNWVGEVLCRATTIVFGNNMYS 179

Qy 200 ILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQP 259
Db 180 ILILTGMGINRYLATAHPTVYQKLPKPSLSLMCGIVVMVFLYMLPFFILKQBYLVHS 239

Qy 260 DIITCHDVHNTCESSPFPQYFISLAPFGFLIPFLVLIICYAAIIRTLNAYDHRWLWY 319
Db 240 EITTCVDWVACESPSPFRFYFVSLAPFGFLIPFLVLIICYAAIIRTLNAYDHRWLWY 299

Qy 320 KASILLIIVIFTCFAPSNIILIIHHANYNNYNTDGLFYVLIACLSGLNSCLDPFLYEL 379
Db 300 KAVULLIIVIFTCFAPNIILIIHHANYNNYNTDGLFYVLIACLSGLNSCLDPFLYFV 359

Qy 380 MSK 382
Db 360 MSK 362
```

```
RESULT 3
PAR3 RAT STANDARD; PRT; 368 AA.
AC Q920EL;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 3 precursor (PAR-3) (Thrombin receptor-
DE like 2) (Coagulation factor II receptor-like 2).
GN F2RL2 OR PAR3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RC SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Spleen;
RA Chien E.K., Marietti S., Mendoza J., Philippe M.;
RA "Cloning of the rat protease activated receptor isoforms 3 and 4.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for activated thrombin coupled to G proteins
CC that stimulate phosphoinositide hydrolysis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PFM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
```

```
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF310076; AAL26789.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G PROTEIN RECP FL1; FALSE_NEG.
DR PROSITE; PS00262; G PROTEIN RECP FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 37 REMOVED FOR RECEPTOR ACTIVATION (BY
FT SIMILARITY).
FT CHAIN 38 368 PROTEINASE ACTIVATED RECEPTOR 3.
FT DOMAIN 38 93 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 94 114 1 (POTENTIAL).
FT DOMAIN 115 123 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 124 144 2 (POTENTIAL).
FT DOMAIN 145 166 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 167 187 3 (POTENTIAL).
FT DOMAIN 188 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 229 4 (POTENTIAL).
FT DOMAIN 230 257 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 258 278 5 (POTENTIAL).
FT DOMAIN 279 300 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 301 321 6 (POTENTIAL).
FT DOMAIN 322 338 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 339 359 7 (POTENTIAL).
FT DOMAIN 360 368 CYTOPLASMIC (POTENTIAL).
FT SITE 37 38 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DISULFID 164 243 BY SIMILARITY.
FT CARBOHYD 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 368 AA; 41795 MW; CE0E94EDA3B80EF1 CRC64;

Query Match 65.6%; Score 1414; DB 1; Length 368;
Best Local Similarity 70.4%; Pred. No. 1.9e-70;
Matches 262; Conservative 41; Mismatches 65; Indels 4; Gaps 2;

Qy 18 IKMKALIFAAAGLLLLPTFCQSGMENDTNNAKPTLPKTRFGAPNSFEFPFSALEG 77
Db 1 MEMKVLILVGRLLFLPTTVCQSGMKRVSDNSA---LTAESFNG-NEHSFEFPPLSDIEG 56

Qy 78 WTGATITVKIKPEESASHLVKNATMGVLTSSLSKLPALYLLVGVGVPANAVTLWLM 137
Db 57 WTGATITVKAKPEESITTLHVNNATMGVLTSSLSKLPALYLLVGVGVPANIVTLWK 116

Qy 138 LFPRTSICITVFTYTNLAIDFLFCVTLPFKIAYHLNGNNWVGEVLCRATTIVFGNNY 197
Db 117 LSSRTKSIISLVFHTNLATDLFCVTLPFKIAYHLNGNDWVGEVLCRATTIVFGNNY 176

Qy 198 CSTILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLV 257
Db 177 CALILTCMGINRYLATVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLV 236

Qy 258 QPDITTCVDVHNTCESSPFPQYFISLAPFGFLIPFLVLIICYAAIIRTLNAYDHRWLW 317
Db 237 QPGITTCVDVHNTCESSPFPQYFISLAPFGFLIPFLVLIICYAAIIRTLNAYDHRWLW 296

Qy 318 YKVASILLIIVIFTCFAPSNIILIIHHANYNNYNTDGLFYVLIACLSGLNSCLDPFLY 377
Db 297 YKAVULLIIVIFTCFAPNIILIIHHANYNNYNTDGLFYVLIACLSGLNSCLDPFLY 356

Qy 378 FLMSKTRNHSSTA 389
Db 357 FIMSKIVDQLTS 368
```



```
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
FT SIGNAL 1 20
FT PROPEP 21 42
FT CHAIN 43 420
FT DOMAIN 43 101
FT TRANSSEM 102 127
FT DOMAIN 128 136
FT TRANSSEM 137 156
FT DOMAIN 157 175
FT TRANSSEM 176 197
FT DOMAIN 198 217
FT TRANSSEM 218 238
FT DOMAIN 239 267
FT TRANSSEM 268 287
FT DOMAIN 288 310
FT TRANSSEM 311 333
FT DOMAIN 334 345
FT TRANSSEM 346 369
FT DOMAIN 370 420
FT TRANSSEM 295 300
FT SITE 42 43
FT DISULFID 174 253
FT CARBOHYD 38 38
FT CARBOHYD 86 86
FT MUTAGEN 42 42
SQ SEQUENCE 420 AA; 47435 MW; D5163F56AFE12372 CRC64;

Query Match 27.8%; Score 599.5; DB 1; Length 420;
Best Local Similarity 34.6%; Pred. No. 2.7e-26;
Matches 128; Conservative 68; Mismatches 149; Indels 25; Gaps 8;

Qy 30 LLLLLPT-----FCQSGMENDTNLAKPLPIKTR--CAPNSPEEPFSALEGW 79
Db 8 LLLLLTLGAGSLCLANSTQAKGHSNNMTIKTRIPDSESEEPFSALEGW 67

Qy 80 GATTIVKICPEESASHLHVKNATMGYLTSSLSKLIPIAYLLVFGVGPANAVTLWMLF 139
Db 68 GSGDQAPVSRARKPIRRNITKEAEQYLSQWLTKFVPSLYTVFVGLPLNLLAIIFL 127

Qy 140 PRTSICITVPEY-TNLAIADFLFCVTLPFKTAYHLNGNWNVFGVGLCAITVIGNYMC 198
Db 128 FKMVKRPVAVYMLNLAIAIDFFVSVLPFKTAYHLSGNDWLFPGMCMCRIVTAIFYCNYC 187

Qy 199 SILLACISINRYLAIVHPFTYRGLPKHYALVTCGLVATVFLYMLPFFTLKQEYLVQ 258
Db 188 SVLLIASISVDRFLAVYPMHLSWRWMSRAYMACSFILWLSIASTIPLLVTEQTKIPR 247

Qy 259 PDITTCCHDVHNTCSSSPFQLYFYFISLAPFGFLIPFVLIYCYAAIINTLNAYD----- 312
Db 248 LDITTCCHDVL-DKDLKDFYIYFSSFCCLFFFPFIITTCYIGIIRSLSSSIENSCK 306

Qy 313 -HRWLWYKASLLIIVFTICFAPSNIILIIHHANYNNYNTDGLFYLIALLCLGSLNSC 371
Db 307 KTRALF---LAVVLCVFIICFGPTNVLFLTH---YLQEAERFLFYAIIACVGSVSC 360

Qy 372 LDPLFLPLMS 381
Db 361 LDPLIYYVAS 370

RESULT 6
PAR2_RAT
ID PAR2_RAT STANDARD; PRT; 397 AA.
AC Q63645;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE like 1) (Coagulation factor II receptor-like 1).
GN F2RL1 OR PAR2.
OS Rattus norvegicus (Rat).
```

```
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN RN SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Intestine, and Kidney;
RA MEDLINE=96358009; PubMed=8762073;
RA Safieddine M., Al-Ani B., Cheng C.H., Wang L., Hollenberg M.D.;
RT "Rat proteinase-activated receptor-2 (PAR-2): cDNA sequence and
RT activity of receptor-derived peptides in gastric and vascular
RT tissue.";
RL Br. J. Pharmacol. 118:521-530(1996).
CC CC -1- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U61373; AAC52703.1; -
CC HSP; P34996; IIDD
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 25
FT PROPEP 26 36
FT CHAIN 37 397
FT DOMAIN 37 75
FT TRANSSEM 76 101
FT DOMAIN 102 110
FT TRANSSEM 111 130
FT DOMAIN 131 149
FT TRANSSEM 150 171
FT DOMAIN 172 190
FT TRANSSEM 191 211
FT DOMAIN 212 241
FT TRANSSEM 242 260
FT DOMAIN 261 285
FT TRANSSEM 286 308
FT DOMAIN 309 323
FT TRANSSEM 324 347
FT DOMAIN 348 397
FT TRANSSEM 383 392
FT SITE 36 37
FT DISULFID 148 226
FT CARBOHYD 23 23
FT CARBOHYD 222 222
SQ SEQUENCE 397 AA; 44440 MW; 27DE0C3ACABC798D CRC64;

Query Match 27.6%; Score 595; DB 1; Length 397;
Best Local Similarity 33.4%; Pred. No. 4.6e-26;
Matches 134; Conservative 71; Mismatches 128; Indels 68; Gaps 12;

Qy 22 ALIFAAAGLLLLLPTFCQSGMENDTNLAKPLPIKTRG---APPNSPEEPFSALEGW 78
Db 5 SLAWLLGGITLL-----AASASCNRTVNPAGPNSKGRSLIGRLDTPP-----PTTG--- 50

Qy 79 TGATITVKIKCPESASHLHVKNATMGYLTSSLSKLIPIAYLLVFGVGPANAVTLWML 138
Db 51 KGAPVEGFGVDFSFAS-----VLTKUTVFLPVIYIIVFVIGLPSGNMLWVF 100
```

```

139 FFRTRISCTTVEY-TNLAIADFLPCVTLPPKIAHYLNGNNVWGEVLCRATTVIFVGNMY 197
101 FFRTKKHPAVIYMANALADLLSVIWFPLKISVHLGNDWTYGDALCKVLIGFFVGNMY 160
198 CSILLACISINRYLAIVHPFTYRGLPKHYVALTCGL---VWATFVLYMLPFILKQEY 254
161 CSILFMTCLSVQRYVWVIVNPMGH---SRKRANAVGVSLAIWLLFLVITPLVYVMEQTI 216
255 YLVPDITTHCHDVHNTCESSPPFOLY-----YFISLAFGFLIPFLVLIYCYAAIIRTL 308
217 YIPALNITTHCHDV-----LPEEVLVGDMPFSYFLSLAIGVFLFPALLTASAYVLMKTL 269
309 NA-----YDRWLMVYKASILLIIVTICFAPSNIILITHANYNNYNTDGLYFIYLIA 362
270 RSSAMDEHSEKRRRAIRLIIIVLUSMYFICFAPSNVLLVHYFLIKSQROSHVVALYIVA 329
363 LCLGLSLNSCLDPFLYFLMSKTRNHSHTAYLTKXXNDLREQGQ 403
330 LCLSTLNSCLDPFYVYVFSK-----DFRDQAR 356

RESULT 7
PAR2 HUMAN
ID PAR2 HUMAN STANDARD; PRT; 397 AA.
AC P55085; Q13317; Q13346;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 2 precursor (PAR-2) (Thrombin receptor-
DE like 1) (Coagulation factor II receptor-like 1).
GN F2RL1 OR PAR2 OR GPR11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96048032; PubMed=7556175;
RA Nystedt S., Emlison K., Larsson A.-K., Stroenbeck B., Sundelin J.;
RT "Molecular cloning and functional expression of the gene encoding the
RT human proteinase-activated receptor 2.";
RL Eur. J. Biochem. 232:84-89(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=96177879; PubMed=8615752;
RA Boehm S.K., Kong W., Broemme D., Smeekens S.P., Anderson D.C.,
RA Connolly A.J., Kahn M.L., Nelken N.A., Coughlin S.R., Payan D.G.,
RA Bunnett N.W.;
RT "Molecular cloning, expression and potential functions of the human
RT proteinase-activated receptor-2.";
RL Biochem. J. 314:1009-1016(1996).
RN [3]
RP SEQUENCE OF 29-397 FROM N.A.
RX MEDLINE=96379236; PubMed=8784787;
RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A.J., Shi Y.P.,
RA Wu R., Lin C.C., Coughlin S.R.;
RT "Conserved structure and adjacent location of the thrombin receptor
RT and protease-activated receptor 2 genes define a protease-activated
RT receptor gene cluster.";
RL Mol. Med. 2:349-357(1996).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS PHE-21; GLN-270 AND ALA-291.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary, and Pancreas;
RA Strauberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for trypsin and trypsin-like enzymes coupled to
CC G proteins that stimulate phosphoinositide hydrolysis. May have a

```

```

role in the regulation of vascular tone.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- TISSUE SPECIFICITY: WIDELY EXPRESSED IN TISSUES WITH ESPECIALLY
HIGH LEVELS IN PANCREAS, LIVER, KIDNEY, SMALL INTESTINE, AND
COLON. MODERATE EXPRESSION IS DETECTED IN MANY ORGANS, BUT NONE IN
BRAIN OR SKELETAL MUSCLE.
-!- PTM: A proteolytic cleavage generates a new amino terminus that
functions as a tethered ligand.
-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; Z49993; CAA90290.1; -.
DR EMBL; Z49994; CAA90290.1; JOINED.
DR EMBL; U34038; AAB47871.1; -.
DR EMBL; U36753; AAA90957.1; -.
DR EMBL; AF400075; AAK77914.1; -.
DR EMBL; BC012453; AAH12453.1; -.
DR EMBL; BC018130; AAH18130.1; -.
DR HSSP; P34996; 1DDO.
DR Genew; HGNC:3538; F2RL1.
DR MIM; 600933; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECP FL 1; FALSE_NEG.
DR PROSITE; PS0282; G-PROTEIN RECP FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Polymorphism.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 36 REMOVED FOR RECEPTOR ACTIVATION
FT CHAIN 37 397 (BY SIMILARITY)
FT DOMAIN 37 75 PROTEINASE ACTIVATED RECEPTOR 2.
FT TRANSMEM 76 101 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 102 110 1 (POTENTIAL).
FT TRANSMEM 111 130 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 131 149 2 (POTENTIAL).
FT TRANSMEM 150 171 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 172 190 3 (POTENTIAL).
FT TRANSMEM 191 211 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 212 241 4 (POTENTIAL).
FT TRANSMEM 242 260 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 261 285 5 (POTENTIAL).
FT TRANSMEM 286 308 6 (POTENTIAL).
FT DOMAIN 309 323 7 (POTENTIAL).
FT TRANSMEM 324 347 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 348 397 7 (POTENTIAL).
FT TRANSMEM 349 397 CYTOPLASMIC (POTENTIAL).
FT SITE 36 37 POLY-SER.
FT DISULFID 148 226 CLEAVAGE (BY TRYPSIN) (POTENTIAL).
FT CARBOHYD 30 30 BY SIMILARITY.
FT CARBOHYD 222 222 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 21 21 S -> F.
FT VARIANT 270 270 /FTID=VAR_012846.
FT VARIANT 291 291 R -> Q.
FT VARIANT 291 291 /FTID=VAR_012847.
FT VARIANT 291 291 T -> A.
FT CONFLICT 138 138 /FTID=VAR_012848.
FT CONFLICT 291 291 G -> A (IN REF. 2).
FT CONFLICT 291 291 T -> S (IN REF. 5; AAH18130).
FT SEQUENCE 397 AA; 44126 MW; 1A4E1D5AB9B362B CRC64;
Query Match 27.3%; Score 589.5; DB 1; Length 397;
Best Local Similarity 35.0%; Pred. No. 9.1e-26;
Matches 137; Conservative 65; Mismatches 138; Indels 51; Gaps 13;

```



```
Qy 28 AGLLLLPTTCQSGWENDTNNAKLPITPKTFRGAPPNSPEEFPSALEGWTGATITVKI 87
Db 12 AAIIILAAASLSC-SGTIOGTRSSKGRSLIGKVDGT-----SHVTG-KGVTIVETVF 59
Qy 88 KCPEESASHLVKNAWTKWYLTSSLSKLIPIAYLLVGVVGVANAVTLWMLFRTSICIT 147
Db 60 SVDEFSAS-----VLTGKUTTVFLPIVYIVFVGVGLPSNGMALWFLFKTKKHP 109
Qy 148 TVFY-TNLAIADFLFCVTLFPKIAHYHLNGNNWVEVLCRATTIVFYGNMYCSILLIACI 206
Db 110 AVIYMANLADLLSVIWFPLKIAHYHNGNNWVEVLCRATTIVFYGNMYCSILLIACI 169
Qy 207 SINRYLAIVHPFTVRGLPKHTYALVTCGLVWATVFLWMLFFILKQBYLVQDPDITCHD 266
Db 170 SVORYVIVNPMGH-SRKANKAIAIGISLAIIWLLVLTIPLYVVKQIFIPALNITTCCHD 228
Qy 267 VHTNCSSSPOLY-----YFISLAFPGELIPVLIYCYAAIIRTLNA-----VDHR 314
Db 229 V-----LPEQLLVGMFNYFLSLAIGVFLPFLTASAYVLMIRLSAMDENSEKK 281
Qy 315 WLWVYKASLLIIVFTICFAPSNIILIIHANYNNYNTDG---LYFYLIACLCIGSLNSC 371
Db 282 RKRAIKLIVTVLAMYLCFTPSNLLLVH---YFLIKSQGSHVYALYIYVALCLSTLNSC 338
Qy 372 LDPELYFLMSKT-RNHS-TAYLTXXNDLRE 400
Db 339 IDPFVYVSHDRDRAKNAIICRSVRTVKQ 369

RESULT 8
PAR4 RAT
ID -PAR4 RAT STANDARD; PRT; 395 AA.
AC O920EO;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE like 3) (Coagulation factor II receptor-like 3).
GN F2RL3 OR PAR4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RC STRAIN=Sprague-Dawley; TISSUE=Duoenum;
RA Hoogerwerf W.A.; Lee-Hallmich H.; Paricha P.J.;
RT "Proteinase activated receptor 4 from rat duodenal library.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN (2)
RC STRAIN=Sprague-Dawley; TISSUE=Myometrium;
RA Chien E.K.; Marietti S.; Mendoza J.; Philippe M.;
RT "Cloning of the rat protease activated receptor isoforms 3 and 4.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC proteins that stimulate phosphoinositide hydrolysis. May play a
CC role in platelets activation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF310216; AAL26790.1; -
DR EMBL; AF269246; AAK58604.2; -
```

```
DR InterPro; IPR000276; GPCR Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 16 POTENTIAL.
FT PROPEP 17 58 REMOVED FOR RECEPTOR ACTIVATION
FT FT (BY SIMILARITY).
FT CHAIN 59 395 PROTEINASE ACTIVATED RECEPTOR 4.
FT DOMAIN 59 93 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 94 114 1 (POTENTIAL).
FT DOMAIN 115 119 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 120 140 2 (POTENTIAL).
FT DOMAIN 141 161 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162 182 3 (POTENTIAL).
FT DOMAIN 183 203 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 204 224 4 (POTENTIAL).
FT DOMAIN 225 254 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 255 275 5 (POTENTIAL).
FT DOMAIN 276 295 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 317 330 6 (POTENTIAL).
FT DOMAIN 331 354 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 355 395 7 (POTENTIAL).
FT SITE 58 59 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 98 101 CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT DOMAIN 373 376 POLY-VAL.
FT DISULFID 160 239 POLY-SER.
FT CONFLICT 1 9 BY SIMILARITY.
FT CONFLICT 192 192 MISSING (IN REF. 2).
FT CONFLICT 192 192 S -> P (IN REF. 2).
SQ SEQUENCE 395 AA; 42943 MW; 43B36D0DA22FAFAC CRC64;

Query Match 24.8%; Score 535; DB 1; Length 395;
Best Local Similarity 37.2%; Pred. No. 8.1e-23;
Matches 109; Conservative 54; Mismatches 128; Indels 2; Gaps 2;

Qy 89 CPBESASHLVKNAWTKWYLTSSLSKLIPIAYLLVGVVGVANAVTLWMLFFRTSICIT 148
Db 65 CANNSDI-LLEPASSEALLGWVPTRLVPAIYGLVGVVGLPANGALWLVATRVPRPST 123
Qy 149 VFTYLAIAIDFLFCVTLFPKIAHYHLNGNNWVEVLCRATTIVFYGNMYCSILLIACISI 208
Db 124 ILLMLAVADLLALVLPRLVHLRQRPFGAEACRVATAALYGHMYGSVLLAAVSL 183
Qy 209 NRVLAIVHPFTVRGLPKHTYALVTCGLVWATVFLWMLFFILKQBYLVQDPDITTCDDVH 268
Db 184 DRYLALVHSURALRQRLTTILCVLWLSAATLVLPFLFHRQTLFLLAGSDRMCLDAL 243
Qy 269 NTCESSPFQLYYFISLAFPGELIPVLIYCYAAIIRTLNAYDHRWLVYKASLLIIVI 328
Db 244 PLAEQTSNHR-PAFICLAVLGCVFPLAMVLCYGATRALAANGQRYSHAVRLTALVLS 302
Qy 329 FTICFAPSNIILIIHANYNNYNTDGLYFYLIACLCIGSLNSCLDPELYFLMS 381
Db 303 AVAAFTPSNVLLVHNSPSPANGNLYGAYVPSLALSTLNSCLDPELYFLMS 355

RESULT 9
PAR4 MOUSE
ID -PAR4 MOUSE STANDARD; PRT; 396 AA.
AC O88634;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE like 3) (Coagulation factor II receptor-like 3).
GN F2RL3 OR PAR4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
```

RP SEQUENCE FROM N.A.
RX MEDLINE=9839762; PubMed=9722561;
RA Kahn M.L., Hammes S.R., Botka C., Coughlin S.R.;
RT "Gene and locus structure and chromosomal localization of the
RT protease-activated receptor gene family.";
J. Biol. Chem. 273:23290-23296 (1998).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=9837991; PubMed=9716134;
RA Kahn M.L., Zheng Y.-W., Huang W., Bigornia V., Zeng D., Moff S.,
RA Farese R.V. Jr., Tam C., Coughlin S.R.;
RT "A dual thrombin receptor system for platelet activation.";
Nature 394:690-694 (1998).
RL Nature 394:690-694 (1998).
CC -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC proteins that stimulate phosphoinositide hydrolysis. May play a
CC role in platelets activation.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in the spleen. Slight
CC expression in the heart, lung, skeletal muscle and kidney. No
CC detectable expression in brain, liver or testis. Also detected in
CC platelets.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF080215; AAC29861.1; -
DR MGD; MGI:1298207; F2r13.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 16
FT PROPEP 17 59
FT CHAIN 60 396
FT DOMAIN 60 94
FT TRANSMEM 95 115
FT DOMAIN 116 120
FT TRANSMEM 121 141
FT DOMAIN 142 162
FT TRANSMEM 163 183
FT DOMAIN 184 203
FT TRANSMEM 204 224
FT DOMAIN 225 235
FT TRANSMEM 236 256
FT DOMAIN 257 277
FT TRANSMEM 278 295
FT DOMAIN 296 316
FT TRANSMEM 317 331
FT TRANSMEM 332 355
FT DOMAIN 356 396
FT DOMAIN 374 377
FT SITE 59 60
FT DISULFID 161 240
FT CARBOHYD 68 68
FT SEQUENCE 396 AA; 42703 MW; E96636F4804FD3E7 CRC64;
Query Match 24.6%; Score 531.5; DB 1; Length 396;
Best Local Similarity 35.7%; Pred. No. 1.3e-22;
Matches 112; Conservative 57; Mismatches 128; Indels 17; Gaps 2;
QY 84 TVKIKCPES-----ASHLHVKNATMGVLTSLSTKLIPALYLVFVVG 127
DB 44 TVELKEPKSDKPNRPGKFCANDSTLELPASSQALLGWSTRLVYGLVAVG 103

QY 128 VPAVAVTLWMLFFRTSRISCTTVFVTLAIADFLFCVTLPFKIAVHLNNGNWFGEVLCRA 187
DB 104 LPANGLALWLAIVPRLPSTILLTNLAVADSLALVPPRLAYHLRQWRPFGEACRV 163
QY 188 TTIVFYGNMYCSILLACISINRYLAIVHPTVYRGLPKHTYALVTCGLVWATVLYMLPF 247
DB 164 ATAALYGHMYGVSLLLAASLDYRLALVHPLRARALRGQRLLTTGLCLVAVLWLSAATLALPL 223
QY 248 FILKQEVYLVQPDITTCCHDVHNTCESSSPQLYYFISLAFFGLIPFVLIYCYAAIIRT 307
DB 224 TLHQTPRLAGSDRMUCHDALPTEQTSRWR-PAFICLAVLGCFVPLAMGLCYGATLRA 282
QY 308 LNAVYDRHVLWYKASLLILVITFCFAPSNIILIIHHANYNNNTDGLFYIYIALCLGS 367
DB 283 LAANGQRYSHALRLTALVLFSAVASFTPSNVLVHLVSNPSPEAGNLYGAYVPSLALST 342
QY 368 LNSCLDPFLYFLMS 381
DB 343 LNSCVDPFYIYVVS 356
RESULT 10
PAR1 MOUSE
ID PAR1_MOUSE STANDARD; PRT; 430 AA.
AC P30558; P97507;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN F2R OR PAR1 OR CF2R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Coughlin S.R.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=96379236; PubMed=8784787;
RA Kahn M.L., Ishii K., Kuo W.L., Piper M., Connolly A., Shi Y.P.,
RA Wu R., Lin C.C., Coughlin S.R.;
RT "Conserved structure and adjacent location of the thrombin receptor
RT and protease-activated receptor 2 genes define a protease-activated
RL receptor gene cluster.";
RL Mol. Med. 2:349-357 (1996).
RN (3)
RP SEQUENCE OF 1-74 FROM N.A.
RC STRAIN=129/SV;
RX MEDLINE=96327649; PubMed=8678993;
RA Xue J., Jenkins N.A., Gilbert D.J., Copeland N.G., Sadler J.E.;
RT "Structure and localization of the thrombin receptor gene on mouse
RT chromosome 13.";
RL Mamm. Genome 7:625-626 (1996).
CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
CC G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```
DR EMBL; L03529; AAA40438.1; -.
DR EMBL; U36757; AAB38308.1; -.
DR EMBL; U36756; AAB38308.1; JOINED.
DR EMBL; U55076; AAB00198.1; -.
DR EMBL; U55075; AAB00198.1; JOINED.
DR HSSP; P34996; 1DDO.
DR MGD; MGI:101802; F2Y.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation.
FT SIGNAL 1 21
FT PROPEP 22 41
FT CHAIN 42 430
FT DOMAIN 42 107
FT TRANSMEM 108 133
FT DOMAIN 134 142
FT TRANSMEM 143 162
FT DOMAIN 163 181
FT TRANSMEM 182 203
FT DOMAIN 204 223
FT TRANSMEM 224 244
FT DOMAIN 245 273
FT TRANSMEM 274 293
FT DOMAIN 294 316
FT TRANSMEM 317 339
FT DOMAIN 340 354
FT TRANSMEM 355 379
FT DOMAIN 380 430
FT DOMAIN 60 65
FT DOMAIN 88 91
FT SITE 41 42
FT DISULFID 180 259
FT CARBOHYD 67 67
FT CARBOHYD 80 80
FT CARBOHYD 255 255
FT CONFLICT 162 162
FT CONFLICT 189 189
FT CONFLICT 223 223
FT CONFLICT 262 262
FT CONFLICT 365 365
SQ SEQUENCE 430 AA; 47790 MW; 395FD64FAE52C9BF CRC64;

Query Match 24.2%; Score 522.5; DB 1; Length 430;
Best Local Similarity 30.8%; Fred. NO. 4.1e-22;
Matches 130; Conservative 82; Mismatches 161; Indels 49; Gaps 16;

QY 21 KALIFAAGLLLLPTF-CQSGMENDTNLAKPTLPKTF--RGAPPNSFEFFPSALEG 77
DB 4 RLLIVALGLSLCGPLLSRVPMQPSERTDATVNPSPFLRNPSENTFELVPLGDEE 63

QY 78 -----WTGATTIVKIKCEESASHLVHVNATWGYLTSSLSKLPALVYLVVGVGP 129
DB 64 EKNESVLLLEGRAVYLNISLPHTPPPPFISEDASGYLTSPWLTFLFMSVYTVIFVLSLP 123

QY 130 AN--AVTLWMLFFRTSICITVYTNLAIDFLCVTLTPFKIAVHLGNVNWVGEVL CRA 187
DB 124 LNVIAIAVFLRMVKK--PAVYMLHLMAVDLVFVSLPKISYFSGTDQFGSGMCRF 182

QY 188 TTVIFYGNMYSIILLACISINRYLAI VHP---FTYRGLPKHTVALVTCGLVWATVFLYM 244
DB 193 ATAIFYGNMYSIMLTVISIDRFVAVYPIQSLSWRTLGRANF---TCVIVWMAIMGV 239

QY 245 LPFFILKQYVLPDIITCHDV--HNTCESSSPQLYFTISLAFPGFLPFLVLIYCYA 302
DB 240 VPLLKEQTRVPGNLITTCVDLSLENMQG---FYSYFSAFSAIFELVPLVIVSTVCT 296

QY 303 AIIRTLN-----AYDRWLWYKVASLLILVIITFCAPSNIILIHANYYY----NN 351
DB 297 SIIRCLSSAVANRSKSRALF---LSAAVFCIFVCFGPTNVLIVH---YLFSLDSPG 350

352 TDGLYFYIYLIALCLGSLNSCLDPPL-YFLMSKTRNHSTAYL-----TKXNDLREQG--P 404
351 TEAAYFAYLLVCVSVSSCCIDPLIYVYASSECORHLVYSILCKESSDPNSCSTGQLMP 410
405 SQ 406
411 SK 412

RESULT 11
PAR4_HUMAN
ID PAR4_HUMAN STANDARD; PRT; 385 AA.
AC Q96RIO; O76067;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Proteinase activated receptor 4 precursor (PAR-4) (Thrombin receptor-
DE like 3) (Coagulation factor II receptor-like 3).
GN F2RL3 OR PAR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389762; PubMed=9722561;
RA Kahn M.L., Hammes S.R., Botka C., Coughlin S.R.;
RT "Gene and locus structure and chromosomal localization of the
RT protease-activated receptor gene family.";
RL J. Biol. Chem. 273:23290-23296(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98379991; PubMed=9716134;
RA Kahn M.L., Zheng Y.-W., Huang W., Bigornia V., Zeng D., Moff S.,
RA Farese R.V. Jr., Tam C., Coughlin S.R.;
RT "A dual thrombin receptor system for platelet activation.";
RL Nature 394:690-694(1998).
RN [3]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-47 AND ARG-68.
RX TISSUE=Lymphoma; PubMed=9618465;
RA Xu W.-F., Andersen H., Whitmore T.E., Presnell S.R., Yee D.P.,
RA Ching A., Gilbert T., Davie E.W., Foster D.C.;
RT "Cloning and characterization of human protease-activated receptor
RT 4.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:6642-6646(1998).
RN [4]
RP SEQUENCE FROM N.A., AND VARIANTS VAL-296 AND LEU-310.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP FUNCTION.
RX MEDLINE=99178892; PubMed=10079109;
RA Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
RA Coughlin S.R.;
RT "Protease-activated receptors 1 and 4 mediate activation of human
RT platelets by thrombin.";
RL J. Clin. Invest. 103:879-887(1999).
CC CC -!- FUNCTION: Receptor for activated thrombin or trypsin coupled to G
CC CC proteins that stimulate phosphoinositide hydrolysis. May play a
CC CC role in platelets activation.
CC CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in lung,
CC CC pancreas, thyroid, testis and small intestine. Not expressed in
CC CC brain, kidney, spinal cord and peripheral blood leukocytes. Also
CC CC detected in platelets.
CC CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
CC CC functions as a tethered ligand.
CC CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```



```
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 425 AA; 47253 MW; 118FCSFB39D4DE2C CRC64;

Query Match
Best Local Similarity 24.1%; Score 520.5; DB 1; Length 425;
Matches 125; Conservative 75; Mismatches 153; Indels 27; Gaps 11;

Qy 21 KALIFAAAGLLLLPTF-CQSGMENDTNNAKPTLPKTKRFGAPPN-SFEFFPFSALEGW 78
Db 4 RRLLLVAACLCGCLGFLSARTRARRPASKATNATLDRSFLLRPNPKYEPFWEDEKNE 63
Qy 79 TGAT-----ITVKICPESASHLVKATMGYLTSSLSKLPAPYLLVFFVGVGPAN--A 132
Db 64 SGLTEYRLVSNKSPLOKLPAPFISDASGYLTSSMLTLFVPSVYTGTVFVSLPVNIMA 123
Qy 133 VTLNMLPRTSICITVFTYNLATADPLFCVTLFPKLYHLNGNWWVGEVLCRATTVIF 192
Db 124 IVFILKMKVKK-PAVYMLHLATADLVFVSFLPKISYLSGSDMDFGSELCRFVTAAP 182
Qy 193 YGNMYCSILLACISINRYLAIVHP---FTYRGLPKHTYALVTCGLVMATVFLYMLPFFI 249
Db 183 YCNMYASILLMTVISIDRLAVYPMQSLSWRTLGRASP---TCLATWALAAGVPLLL 239
Qy 250 LKQYLYVQPDITTDHVDHNTCESSPPQLYYFISLAPFGFLIPFVLIYYCYAAIIRTLN 309
Db 240 KEQTIQVPLNITTDHVDHNTLLEG-YYAYYFSAFSAVFVFLIISTVCVYSIIRCLS 298
Qy 310 -----AYDRWLWYWKASLLIIVFTICFAPSNIILIHANY-YNNNTDGLYFYLI 361
Db 299 SSTVANRKKSRALF---LSAAVFCIFICFGPTNILLIAHYSFLSHTSTTEAAYFAYLL 355
Qy 362 ALCLGSLNSCLDPLFYFLMS 381
Db 356 CVCVSSISCCDPLIYYAS 375

RESULT 13
PARI HUMAN STANDARD; PRT; 425 AA.
AC P25116; Q96RF7; Q9BUN4;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PARI-1) (Thrombin receptor)
DE (Coagulation factor II receptor).
CN F2R OR PARI OR TR OR CP2R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91168254; PubMed=1672265;
RA Vu T.-K.H., Hung D.T., Wheaton V.I., Coughlin S.R.;
RT "Molecular cloning of a functional thrombin receptor reveals a novel
RL proteolytic mechanism of receptor activation.";
RL Cell 64:1057-1068(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
RA Nickerson D.A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RX MEDLINE=99178892; PubMed=10079109;
RA Kahn M.L., Nakanishi-Matsui M., Shapiro M.J., Ishihara H.,
RA Coughlin S.R.;
```

```
RT RT "Protease-activated receptors 1 and 4 mediate activation of human
RL platelets by thrombin.";
CC J. Clin. Invest. 103:879-887(1999).
CC -I- FUNCTION: High affinity receptor for activated thrombin coupled to
CC G proteins that stimulate phosphoinositide hydrolysis. May play a
CC role in platelets activation and in vascular development.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- TISSUE SPECIFICITY: PLATELETS AND VASCULAR ENDOTHELIAL CELLS.
CC -I- PFM: A proteolytic cleavage generates a new amino terminus that
CC functions as a tethered ligand.
CC -I- PFM: Phosphorylated; probably mediating desensitization prior to
CC the uncoupling and internalization of the receptor.
CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62424; AAA36743.1; -.
DR EMBL; AF391809; AA69768.1; -.
DR EMBL; BC002464; AA02464.1; -.
DR PIR; A37912; A37912.
DR HSP; P34996; IDDD.
DR Genew; HGNC:3537; F2R.
DR MIM; 187930; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Blood coagulation; Phosphorylation.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 41 REMOVED FOR RECEPTOR ACTIVATION.
FT CHAIN 42 425 PROTEINASE ACTIVATED RECEPTOR 1.
FT DOMAIN 42 102 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 103 128 1 (POTENTIAL).
FT TRANSNEM 129 137 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 138 157 2 (POTENTIAL).
FT TRANSNEM 158 176 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 177 198 3 (POTENTIAL).
FT TRANSNEM 199 218 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 219 239 4 (POTENTIAL).
FT TRANSNEM 240 268 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 269 288 5 (POTENTIAL).
FT TRANSNEM 289 311 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 312 334 6 (POTENTIAL).
FT TRANSNEM 335 350 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 351 374 7 (POTENTIAL).
FT TRANSNEM 375 425 CYTOPLASMIC (POTENTIAL).
FT SITE 41 42 ASP/GLU-RICH (ACIDIC).
FT SITE 41 42 CLEAVAGE (BY THROMBIN).
FT DISULFID 175 254 BY SIMILARITY.
FT CARBOHYD 35 35 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 238 238 L -> V (IN REF. 1).
FT CONFLICT 335 335 A -> V (IN REF. 3).
FT CONFLICT 364 364 C -> S (IN REF. 1).
SQ SEQUENCE 425 AA; 47440 MW; 41B742A99EEC96AB CRC64;

Query Match
Best Local Similarity 24.1%; Score 519; DB 1; Length 425;
Matches 126; Conservative 76; Mismatches 154; Indels 30; Gaps 11;

Qy 14 GTQVIRKALIFAAAGLLLLPTLPKTKRFGAPPN-SFEFFP 72
```

Db 2 GPRELLVAACFSLCGPLLSARTREARPEKATN---ATLDPRESFLLRNPNDKYEPEWE 57
Qy 73 SALEGWTGAT-----ITVKIKCPESASHLVKATMGYLTSSLSSTKLIPIAYILLVFFVGV 128
Db 58 DEEKNEGLTEYRLVSKNSPLOKQOLPAFISEDASGYLTSSWLTLPVPSVYTGVPVVS 117
Qy 129 PAN--AVTLNMLPRTSICITVYTNLATADFLCVTLFPKAYHLNGNNWVGEVLCR 186
Db 118 PLNMAIVFLLKMKVK--PAVVTMLLATADVLVSVLPFKISYIFSGSDWQFGSEL 176
Qy 187 ATTIVIFGNMVCISILLACISINRYLAIVHP---FTYRGLPKHTYALVTGCLVMAVFLY 243
Db 177 FVTAAFYCNMYASILLMTVISIDRELAIVVPMQSLSWRTLGRASF---TCLAIWALAIA 233
Qy 244 MLPFFILKOEYLVQPDITTCCHDVHNTCESSPPQLYYFISLAPFGFLIPFVLIICYAA 303
Db 234 WPLLKKEQITQVGLNITTCCHDVHNTLEGL-YYAYFFSAFSAVFFVFLIISTVCYVS 292
Qy 304 IIRTLN-----AYDRMLVWYKASLLIIVFTICFAPSNIILIIHNAV-YYNNTDGL 355
Db 293 IIRCLSSAVNRSKKSRALF---LSAAVFCIFICFPTNVLLIAHYSLSHSTTEAA 349
Qy 356 YFIYIALLCGLSCLNDPFLYFLMS 381
Db 350 YFAYLLVCVSVSSICCDPLIYYVAS 375

RESULT 14

PAR1 RAT ID PAR1 RAT STANDARD; PRT; 432 AA.
AC P26824;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proteinase activated receptor 1 precursor (PAR-1) (Thrombin receptor).
GN F2R OR PAR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RX TISSUE=Aortic smooth muscle;
RX MEDLINE=92381002; PubMed=1324917;
RA Zhong C., Hayzer D.J., Corson M.A., Wick K., Runge M.S.;
RT "Molecular cloning of the rat vascular smooth muscle thrombin
receptor. Evidence for in vitro regulation by basic fibroblast growth
factor.";
RL J. Biol. Chem. 267:16975-16979(1992).
CC -!- FUNCTION: High affinity receptor for activated thrombin coupled to
G proteins that stimulate phosphoinositide hydrolysis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- PTM: A proteolytic cleavage generates a new amino terminus that
functions as a tethered ligand.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M81642; AAA42274.1; -;
DR PIR; A43448; A43448.
DR HSSP; P34996; 1DDO.
DR InterPro; IPR000276; GPCR_Rhodops.
DR Pfam; PF00001; 7cml_1;
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;

KW Blood coagulation. 21
FT SIGNAL 1
FT PROPEP 22 45
FT
FT POTENTIAL.
FT REMOVED FOR RECEPTOR ACTIVATION (BY
FT SIMILARITY).
FT PROTEINASE ACTIVATED RECEPTOR 1.
FT EXTRACELLULAR (POTENTIAL);
FT 1 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 2 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 3 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 4 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 5 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT 6 (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).
FT 7 (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT ASP/GLU-RICH (ACIDIC).
FT POLY-PRO.
FT CLEAVAGE (BY THROMBIN) (BY SIMILARITY).
FT BY SIMILARITY.
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 69 69
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 82 82
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 257 257
FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 266 266
FT N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 432 AA; 48280 MW; DD032B97ABA4605 CRC64;
Query Match . 23.4%; Score 505.5; DB 1; Length 432;
Best Local Similarity 31.2%; Pred. No. 3.4e-21;
Matches 134; Conservative 71; Mismatches 163; Indels 61; Gaps 15;
Qy 21 KALIFAAGLLLLPTFCQSGMENDTNNAKPLPIKT-----PRGAPNPFEPFPPS 73
Db 4 RRELLVAVGLSLCGLPSSRVPWRQPESEMYATPVATPNRFFFLRNPSDETFEQPLG 63
Qy 74 -----ALEGWTGATITVKIKCP-----ESASHLVKNATMGYLTSSLSKLI 116
Db 64 DEEKNESIPLEG--RAVYLNKSRFPMPPPPFISEDAS-----GYLTSPWLTFLI 112
Qy 117 PAIYLLVFFVGV PANAVTLNMLFPRTSICITVYF-TNLAIADFLFCVTLFPFIAYHLNG 175
Db 113 PSYTYTFVIVSLPLNLAIAVFRMKVKPAVVMHLAMADVLFVSVLPFKISYIFSG 172
Qy 176 NNWVGEVLCRAITVIFGNMVCISILLACISINRYLAIVHP---FTYRGLPKHTYALVT 232
Db 173 TDWQFGSGMCRFATAACYCNCMYASIMLTMTVISIDRELAIVVPIQSLSWRTLGRANF--T 229
Qy 233 CGLVWATVFLYMLPFFILKOEYLVQPDITTCCHDVHNTCESSPPQLYYFISLAPFGFLI 292
Db 230 CVVIMWNAIMGVVPVLLKKEQTQVPGNITTCCHDVHNTLEGL-FYSYIFSAFSAIFFLV 288
Qy 293 PFVLIIVCYAAIIRTLN-----AYDRMLVWYKASLLIIVFTICFAPSNIILIIHNA 345
Db 289 PLIISIVCYTISIRCLSSAVNRSKKSRALF---LSAAVFCIFVFCGNTVLLIIVYL 345
Qy 346 NYIYN-NTDGLYFIYIALLCGLSCLNDPFL-YFLMSKTRNHSYAL-----TKXNDLR 399
Db 346 LLSDSPGTETAYFAYLLCVCTSVASCIDPLIYVYASSECQKHLYSILCCRESSDSNCSN 405
Qy 400 EQQ--PSQ 406
Db 406 STGQLMPSK 414

RESULT 15

PAR1_CRIL0
ID PAR1_CRIL0 STANDARD; PRT; 428 AA.
AC Q00991; Q60461;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X99953; CAA68213.1; -
HSP; P34996; 1DDD.
InterPro; IPR000276; GPCR_Rhodpsn.
PRINTS; PR00237; GPCR_Rhodopsin.
PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein.
DOMAIN 1 49
TRANSMEM 50 70
DOMAIN 71 79
TRANSMEM 80 100
DOMAIN 101 118
TRANSMEM 119 139
DOMAIN 140 161
TRANSMEM 162 182
DOMAIN 183 210
TRANSMEM 211 231
DOMAIN 232 254
TRANSMEM 255 275
DOMAIN 276 292
TRANSMEM 293 316
DOMAIN 317 537
TRANSMEM 538 557
DISULFID 116 193
CARBOHYD 26 29
SEQUENCE 537 AA; 62024 MW; B2CF24812F3C19F2 CRC64;

Query Match 19.8%; Score 428; DB 1; Length 537;

Best Local Similarity 36.0%; Pred. No. 6.4e-17;
Matches 105; Conservative 41; Mismatches 102; Indels 44; Gaps 12;

Qy 115 LIPAIYLLVFGVGPANAVTLWMLFFRTS-ICTVFVTLAIADFLCVTLRPKIAVHL 173
Db 45 LIPVSYSAVFGVGPANAVTLWMLFFRTS-ICTVFVTLAIADFLCVTLRPKIAVHL 173
Qy 174 NGNNVFGVLCRATTVIFVGNMYSILLACISINRYLAIVHFT--YRGLPKHYALV 231
Db 105 DKNWPFGEVLCVRLFLFYANLYSSILFLTCISVHRYGVCHFTSILRRMNAKHAY-V 162
Qy 232 TCGLWATVFLYMLPPFLKQEVLYVQDDI--TTCVDVHTCESSPPQLYFISLAPFG 289
Db 163 ICALVWLSVTLCLVFNLI--FVTVSPKVKNTICHD-TRPEDFARYVEYSTAIC 214
Qy 290 FL--IPFVLLIYCYAAIIR-----TLNAYDHRMLWYVKASILLIVIFTICFAP 335
Db 215 LFGIPCLITAGCYGLMTRMLKPIVSGNQOTLPSYKR---SIKTIIFVMAIFALCFMP 271
Qy 336 SNIIILIHANYNNNTDGL--YFI-----YLIALCIGLSLNSCLDPLFLVL 379
Db 272 PHIRTEL---YYARLLGKICVYALNVINVTYKTRPLASANSCLDPLFLVL 319

RESULT 17

P2YR MELGA

ID P2YR MELGA STANDARD; PRT; 362 AA.

AC P49652;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor) (6H1

DE orphan receptor).

GN P2Y1.

OS Meleagris gallopavo (Common turkey).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.

OC NCBI_TaxID=9103;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;
RX MEDLINE=94335907; PubMed=8058061;
RA Filtz T.M., Li Q., Boyer J.L., Nicholas R.A., Harden T.K.;
RT "Expression of a cloned P2Y purinergic receptor that couples to
RT phospholipase C";
RT Mol. Pharmacol. 46:8-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97392456; PubMed=9240460;
RA Li Q., Schachter J.B., Harden T.K., Nicholas R.A.;
RT "The 6H1 orphan receptor, claimed to be the P2Y5 receptor, does not
RT mediate nucleotide-promoted second messenger responses";
RT Biochem. Biophys. Res. Commun. 236:455-460(1997).
CC -1- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
CC INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: MAINLY FOUND IN BLOOD, BRAIN, AND LUNG. TO A
CC LESSER EXTENT IN STOMACH, GUT, AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; U09842; AAA18784.1; -

EMBL; AF012103; AAB65428.1; -

HSP; P34996; 1DDD.

InterPro; IPR000276; GPCR_Rhodpsn.

Pfam; PF00001; 7tm 1; 1.

PRINTS; PR00237; GPCR_Rhodopsin.

PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.

PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.

G-protein coupled receptor; Transmembrane; Glycoprotein.

DOMAIN 1 41
TRANSMEM 42 63
DOMAIN 64 76
TRANSMEM 77 98
DOMAIN 99 115
TRANSMEM 116 136
DOMAIN 137 155
TRANSMEM 156 177
DOMAIN 178 207
TRANSMEM 208 227
DOMAIN 228 254
TRANSMEM 255 274
DOMAIN 275 292
TRANSMEM 293 317
DOMAIN 318 362
DISULFID 113 191
CARBOHYD 11 11
CARBOHYD 26 26
CARBOHYD 102 102
CARBOHYD 186 186
SEQUENCE 362 AA; 41180 MW; 3E128AB9EB64349C CRC64;

Query Match 19.7%; Score 426; DB 1; Length 362;

Best Local Similarity 30.4%; Pred. No. 6.1e-17;
Matches 101; Conservative 59; Mismatches 132; Indels 40; Gaps 9;

Qy 105 GYLTSLSLK-----LIPAIYLLVFGVGPANAVTLWMLFFRTSIC-TTVFT 152

Db 21 GWAAGNASTKSLTKTGQFYLYLFTVILVFTGLGNSVAIWMFVHMRPWSGISVYMF 80

Qy 153 NLAIADFLCVTLPPFKIAYHLNGNNVFGVLCRATTVIFVGNMYSILLACISINRYL 212

Db 81 NLALADFLYVLTLPALIFVFNKTDWFGDMCKLQRFIFHVNLYGSILFTCTSVHRYT 140

QY 213 AIVHPFYRGLPKHYALTCLVWATVFLVWLPFFILKQBY--LVQPDITTCDDVH 268
 Db 141 GWHPLSLGLRKLKKNVYSSLVWALVAVIAPIL-----FSGTGVRNKTITCYD-- 193
 QY 269 NTCSSPPQLYYFISLAFFGLIPFVLIIYCYAAIIRTLNAYD-----HRLWLVYKA 321
 Db 194 TTADEYLRSYFVSMCTVFNFCIPFVILGCGYGLIKVLYKOLDNSPLRKSIIYV-- 251
 QY 322 SLILVITFTICFAPSNIILIIH-HANYYYNN-----TDGLYFYIYALCIGLSNLSCLDP 374
 Db 252 -IIVLVFAVSGLPFHVMKTLNRLARLDFQPMCAFNKQVYATYQVTRGLASLNSCVD 310
 QY 375 FLYELMSKTRNHSYATLTKXNDLREQQPSQ 406
 Db 311 ILYFLAGDTFRRLSRATRKSSRRSEPNVQSK 342

RESULT 18
 P2YR_CHICK
 ID P2YR_CHICK STANDARD; PRT; 362 AA.
 AC P34996;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
 GN P2Y1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93285340; PubMed=8508924;
 RA Webb T.E., Simon J., Krishek B.J., Bateson A.N., Smart T.G.,
 RA King B.F., Burnstock G., Barnard E.A.;
 RT "Cloning and functional expression of a brain G-protein-coupled ATP
 RT receptor.";
 RL FEBS Lett. 324:219-225 (1993).
 RN (2)
 RP 3D-STRUCTURE MODELING.
 RX MEDLINE=97026278; PubMed=8872457;
 RA van Rhee A.M., Fischer B., van Galen P.J.M., Jacobson K.A.;
 RT "Modelling the p2Y purinoceptor using rhodopsin as template.";
 RL Drug Des. Discov. 13:133-140 (1995).
 CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. SEEMS TO MEDIATE ITS ACTION VIA A PERTUSSIS TOXIN
 CC INSENSITIVE G-PROTEIN, PROBABLY BELONGING TO THE GQ FAMILY THAT
 CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: BRAIN, SPINAL CORD, GASTROINTESTINAL TRACT,
 CC SPLEEN AND LEG MUSCLE. IS NOT DETECTED IN THE HEART, LIVER,
 CC STOMACH, LUNG AND KIDNEY.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL databank
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: X73268; CAA51716.1; -
 DR PIR: S33733; S33733.
 DR PDB: 1DDD; 1I-JUL-96.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PR00237; GPCRHHODOPSIN.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECEPTOR F2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; 3D-structure.

FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 42 63 1 (POTENTIAL).
 FT DOMAIN 64 76 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 77 98 2 (POTENTIAL).
 FT DOMAIN 99 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 136 3 (POTENTIAL).
 FT DOMAIN 137 155 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 156 177 4 (POTENTIAL).
 FT DOMAIN 178 207 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 208 227 5 (POTENTIAL).
 FT DOMAIN 228 254 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 255 274 6 (POTENTIAL).
 FT DOMAIN 275 292 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 293 317 7 (POTENTIAL).
 FT DOMAIN 318 362 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 113 191 BY SIMILARITY.
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 362 AA; 41194 MW; A806C88FB9514761 CRC64;
 Query Match 19.7%; Score 424; DB 1; Length 362;
 Best Local Similarity 28.9%; Pred. No. 7.8e-17;
 Matches 101; Conservative 59; Mismatches 143; Indels 46; Gaps 9;
 QY 77 GWTCATITVKKIKPEESASHLVHVNATMGYLTSSLSKLPAPAIYLLVFWVGVSPANAVTLW 136
 Db 21 GWAAGNATTKCSLTGTFQFY-----LPTVYLVITGFLGNSVAVI 63
 QY 137 MLFPRTRISIC-TTVFYTNLAIADFLCVTLPPFKIAYHLNGNNVWVGLVCRATTVIYGN 195
 Db 64 MFVPHMRPWSGISVYMENALADFLYVLTLPALIFYFYFNKTDWIFGDMCKLQRFIFHN 123
 QY 196 MYCSILLACISINRXLAIHVHPTFYRGLPKHYALTCLVWATVFLVWLPFFILKQBY 255
 Db 124 LYSGILFTLCISVHRVTGVVHPLKSLGRLKKNAVSSLVWALVAVIAPIL-----FY 178
 QY 256 ---LVQPDITTCDDVHNTCESSPFQLYYFISLAFFGLIPFVLIIYCYAAIIRTLNAY 311
 Db 179 SGTGVRNKTITCYD--TTADEYLRSYFVSMCTVFNFCIPFVILGCGYGLIKVLYK 236
 QY 312 D-----HRLWLVYKASLLILVITFCFAPSNIILIIH-HANYYYNN-----TDGLYF 357
 Db 237 DLDNSPLRRKSIYLV---IIVLVFAVSGLPFHVMKTLNRLARLDFQPMCAFNKQVYA 293
 QY 358 IYIYALCIGLSNLSCLDPFLYFLMSKTRNHSYATLTKXNDLREQQPSQ 406
 Db 294 TYQVTRGLASLNSCVDPIYFLAGDTFRRLSRATRKSSRRSEPNVQSK 342

RESULT 19
 EB12_HUMAN
 ID EB12_HUMAN STANDARD; PRT; 361 AA.
 AC P32249;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE EBV-induced G protein-coupled receptor 2 (EBI2).
 GN EBI2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93188173; PubMed=838238;
 RA Birkenbach M.P., Josefson K., Yalamanchili R.R., Lenoir G.M.,
 RA Kieff E.;
 RT "Epstein-Barr virus-induced genes: first lymphocyte-specific G
 RT protein-coupled peptide receptors.";
 RL J. Virol. 67:2209-2220 (1993).
 CC -!- FUNCTION: ORPHAN RECEPTOR. PROBABLE MEDIATOR OF EBV EFFECTS ON B

CC LYMPHOCYTES OR OF NORMAL LYMPHOCYTE FUNCTIONS.
CC -1- TISSUE SPECIFICITY: B-LYMPHOCYTE CELL LINES.
CC -1- INDUCTION: BY EBV.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08177; AAA35924.1; --
DR PIR; B45680; B45680.
DR Genew; HGNC:3128; EB12.
DR MIM; 605741; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein.
DR DOMAIN 1 31 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 32 57 1 (POTENTIAL).
FT DOMAIN 58 77 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 78 95 2 (POTENTIAL).
FT DOMAIN 96 105 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 106 127 3 (POTENTIAL).
FT DOMAIN 128 149 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 150 168 4 (POTENTIAL).
FT DOMAIN 169 192 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 193 215 5 (POTENTIAL).
FT DOMAIN 216 241 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 242 265 6 (POTENTIAL).
FT DOMAIN 266 287 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 288 312 7 (POTENTIAL).
FT DOMAIN 313 361 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 126 134 INTERACTION WITH G PROTEINS.
FT DISULFID 104 181 BY SIMILARITY.
SQ SEQUENCE 361 AA; 41224 MW; B5A2171F34C9C67B CRC64;

Query Match 19.0%; Score 410.5; DB 1; Length 361;
Best Local Similarity 30.2%; Pred. No. 4.2e-16;
Matches 99; Conservative 58; Mismatches 114; Indels 57; Gaps 10;

Qy 106 YLTSSLSKTLIPALYLLVFGVGPANAVTLMLEFRTSI-CTTVFYTNLAIDFLFCVT 164
Db 24 YAHSTARIYVPLHYSLVFIIGLVNLLAVVIVQNRKKNSTLYSLNVLISDILFTA 83

Qy 165 LPFKIAYHLNNGVVEVLVCRATVIFYGNMYSICILLACISINRYLAIVHPFTYRGLP 224
Db 84 LPTRIAVYAMGDFNRIGDLCALRTALVFYINTYAGVNFMTCLSIDRFIAVHPLRYNKK 143

Qy 225 KHTVALVTCGLAVATVFLYMLPFTI---LKEYYLVQPDITTCVDVNTCESSS-PFOLY 280
Db 144 RIEHAKGVCIIFWILVFAQLPLLINPMKOE-----AERTICMEYFNFEKSLPWL- 197

Qy 281 YFISLAFEGFLIPFLVLIYCYAAII-----RTLNAYDHRWLWYVK 320
Db 198 --LGACFGYGVLPVLIILICYSQICKLFRYAKONPLTEKSGVKNKALNTI----- 246

Qy 321 ASLILVFTCFAPSNIILIH-----HANYNNYNTDGLYFYL-IALCGLSGLNSCLD 373
Db 247 --ILIVVFLVLCFTPYHVAIIQHKIKLRFNSFLECSQRHSFQISLHFTVCLMNFNCMD 304

Qy 374 PFLYELMSKTRNHSTAYLTXXNDLREQ 401
Db 305 PFIYFFACK-----GYRKRVMRLKQ 326

RESULT 20

GP17 HUMAN STANDARD; PRT; 367 AA.
ID GP17 HUMAN Q9UDZ6; Q9UE21;
AC Q13364; 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable P2Y purinoceptor GP17 (P2Y-like receptor) (R12).
GN GP17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=96145150; PubMed=8558062;
RA Raport C.J., Schweickart V.L., Chantry D., Eddy R.L. Jr., Shows T.B.,
RA Godiska R., Gray P.W.;
RT "New members of the chemokine receptor gene family.";
RL J. Leukoc. Biol. 59:18-23(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Hippocampus;
RC MEDLINE=98181695; PubMed=9523551;
RA Blaesus R.H., Weber R.G., Lichter P., Ogilvie A.;
RT "A novel orphan G protein-coupled receptor primarily expressed in the
RT brain is localized on human chromosomal band 2q21.";
RL J. Neurochem. 70:1357-1365(1998).
CC -1- FUNCTION: Putative receptor for purines coupled to G-proteins (By
CC similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: PRIMARILY EXPRESSED IN BRAIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U33447; AAB16746.1; --
DR EMBL; Y12546; CAA73144.1; --
DR EMBL; Z94154; CAB08107.1; --
DR EMBL; Z94155; CAB08108.1; --
DR HSSP; P34996; IDDD.
DR Genew; HGNC:4471; GP17.
DR MIM; 603071; --
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 64 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 65 85 1 (POTENTIAL).
FT DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 93 113 2 (POTENTIAL).
FT DOMAIN 114 133 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 134 154 3 (POTENTIAL).
FT DOMAIN 155 175 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 176 196 4 (POTENTIAL).
FT DOMAIN 197 223 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 224 244 5 (POTENTIAL).
FT DOMAIN 245 260 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 261 281 6 (POTENTIAL).
FT DOMAIN 282 308 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 309 329 7 (POTENTIAL).
FT DOMAIN 330 367 CYTOPLASMIC (POTENTIAL).
FT DISULFID 132 209 BY SIMILARITY.

OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=T-cell;
 RX MEDLINE=93329058; PubMed=8393036;
 RA Kaplan M.H., Smith D.I., Sundick R.S.;
 RT "Identification of a G protein coupled receptor induced in activated
 T cells";
 RL J. Immunol. 151:628-636(1993).
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- TISSUE SPECIFICITY: INDUCED IN ACTIVATED T-CELLS.
 CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L06109; AAB06587.1; -
 DR HSSP; P34996; 1DDD.
 DR InterPro; IPR00276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
 KW Palmitate.
 FT DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 17 43
 FT DOMAIN 44 52 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 53 76
 FT DOMAIN 77 89 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 90 109
 FT DOMAIN 110 130 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 131 151
 FT DOMAIN 152 178
 FT TRANSMEM 179 206
 FT DOMAIN 207 224 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 225 250
 FT DOMAIN 251 269 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 270 289
 FT DOMAIN 290 308
 FT CARBOHYD 5 5
 FT LIPID 281 281 N-LINKED (GLCNAC...?) (POTENTIAL).
 FT DISULFID 86 165 PALMITATE (BY SIMILARITY).
 SQ SEQUENCE 308 AA; 35597 MW; 4214E969633B6F7D CRC64;
 Query Match 18.8%; Score 406.5; DB 1; Length 308;
 Best Local Similarity 33.5%; Pred. No. 6.1e-16;
 Matches 108; Conservative 52; Mismatches 127; Indels 35; Gaps 11;
 QY 107 LTSSLSK- - - - -LIPAIYLLVGVGVPNAVTLMLFF--RTRSICTTVFYTNLAID 158
 DB 2 VSSNCSTEDSKFYLYGCVSNVFLGLIANCAIYFTFLKVRNE-TTYMLNLAISD 60
 QY 159 FLFCVTLFPFKIAYHLNGNNWVFGVLCRAITTVIFGYNNYCISILLACISINRYLAIYVHPF 218
 DB 61 LLFVFTLPFRI-YFVVRNWFPGVGLCKISVTLFTVNNYSGILFULTCISVDRLAIYVHPF 119
 QY 219 TYRGLPKIYALVTCGLVWATVFLYMLPFFLKQEYLYVQPDITTDHVDHTCESS-SPF 277
 DB 120 RSKLTRKRNARIVCAVWITVLGASTPASPFSQSTNRQNNTEQRTCFE--NFPPESTWKTY 177
 QY 278 QLYYFISLAFPGFLPFLVLIYCYAAIIRTLN---AYDHRWLWVVKASLLI---LVIFTI 331
 DB 178 LSRIVIFIEVGFPIPLINVTCTWVRLTNKPLTSRNLKSKKVLKMFVHLVIFCP 237
 QY 332 CFAPSNILIIH---HANYNNNT--DGLYFIYLIALCLGSLNSCLDPLFYFLMSKTRNH 386
 DB 238 CFVPVNTILYSLMRQTWNCVSVTAVRTWYPTVLCIAVSNCCFDPVIVYFTSDTN-- 295

QY 387 STAYLTXXNDLREQGQPSORT 408
 DB 296 -----SELDKKQVHQNT 308
 RESULT 23
 P2YR HUMAN
 ID P2YR_HUMAN STANDARD; PRT; 373 AA.
 AC P47900;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 1 (ATP receptor) (p2Y1) (Purinergic receptor).
 GN P2RY1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96257237; PubMed=8666290;
 RA Leon C., Vial C., Cazenave J.-P., Gachet C.;
 RT "Cloning and sequencing of a human cDNA encoding endothelial P2Y1
 purinoceptor.";
 RL Gene 171:295-297(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96158962; PubMed=8579591;
 RA Ayyanathan K., Tania W., Harbansjit S., Raghbir A.S., Barnard E.A.,
 RA Kunapuli S.P.;
 RT "Cloning and chromosomal localization of the human P2Y1
 purinoceptor.";
 RL Biochem. Biophys. Res. Commun. 218:783-788(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96203320; PubMed=8630005;
 RA Janssens R., Communi D., Pirotton S., Samson M., Parmentier M.,
 RA Boeynaems J.M.;
 RT "Cloning and tissue distribution of the human P2Y1 receptor.";
 RL Biochem. Biophys. Res. Commun. 221:588-593(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Leon C., Vial C., Weber J., Cazenave J.-P., Gachet C.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 95-373 FROM N.A., FUNCTION, AND INHIBITION STUDIES.
 RC TISSUE=Platelet;
 RX MEDLINE=98113162; PubMed=9442040;
 RA Jin J., Daniel J.L., Kunapuli S.P.;
 RT "Molecular basis for ADP-induced platelet activation. II. The P2Y1
 receptor mediates ADP-induced intracellular calcium mobilization and
 shape change in platelets.";
 RL J. Biol. Chem. 273:2030-2034(1998).
 CC -|- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
 CC ATP AND ADP. IN PLATELETS BINDING TO ADP LEADS TO MOBILIZATION OF
 CC INTRACELLULAR CALCIUM IONS VIA ACTIVATION OF PHOSPHOLIPASE C, A
 CC CHANGE IN PLATELET SHAPE, AND PROBABLY TO PLATELET AGGREGATION.
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -|- INDUCTION: REPRESSED BY THE P2Y1 RECEPTOR-SPECIFIC ANTAGONISTS
 CC A3P5PS, A3P5P AND A2P5P. THESE INHIBIT CALCIUM ION MOBILIZATION
 CC AND SHAPE CHANGE IN PLATELETS.
 CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```
-----
CC EMBL; Z49205; CAA89066.1; -
DR EMBL; U42030; AAA97873.1; -
DR EMBL; U42029; AAA97872.1; -
DR EMBL; S81950; AAB47091.1; -
DR EMBL; AJ006945; CAA07339.1; -
DR EMBL; AF018284; AAB94536.1; -
DR HSSP; P34996; 1DDO.
DR Min; HGNC:8539; P2RY1.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP FL 1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Platelet;
KW Blood coagulation.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 74
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 127 147
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 188
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 219 238
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 266 285
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 328
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 202 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 138 138 MISSING (IN REF. 1).
SQ SEQUENCE 373 AA; 42071 MW; 4DC7C668B4145392 CRC64;
-----
Query Match 18.5%; Score 399; DB 1; Length 373;
Best Local Similarity 29.7%; Pred. No. 1.8e-15;
Matches 113; Conservative 66; Mismatches 143; Indels 58; Gaps 14;
Qy 52 PTLPIKT---PRGAPNSFEFFPSALEGWTGATITVKIKPESASHLHVKNATMGLT 108
Db 7 PAVNGTDAFLAGPSS-----WGNSTV-----ASTAAVSSSFKALT 45
Qy 109 -SSLSUTKLIPAIYLLVFGVPANAVTLMLFFRTRSICTTVFYTNLAIDFLFCVTLTP 166
Db 46 KTGQFYVLPVAVYILVFIIGLNSVAIMVFMFKPWSGISVYMFNLALADFLYVLTLP 105
Qy 167 FKIAVHLNGNNVGEVLCAITTVIYGNMYCSILLACISINYLAIHVHPTVYRGLPKH 226
Db 106 ALIFYYFNKTDFGDMCKLQRFHFVNLVGSILFTCSIAHRYSGVYVPLKSLRLKK 165
Qy 227 TYALVTCGLWATVFLYMLPFILKQEVYLVQPDIT-TCHDVHNTCSSSPFQLYYFI-- 283
Db 166 KNAICISVLWLVJVAISP--ILFYSGTGVKKNKTICTYD-----TSDEYLSRYFIYS 218
Qy 284 ---SLAFGFLIPVLIYCYAAIIRTLNAYD-----HRLWVYKASLLILVIFTICF 333
Db 219 MCTTVAM--FCVPLVLILGCVGLVRLAIYKDLNSPLRRKSIYLV---IIVLVFAVS 273
Qy 334 APSNIIILIH-HANYNN-----TDGLYFIYLIACLSLNSCLDPFLYLMKSTRNH 386
Db 274 IPFHVMTNMLRDLDFQTPAMCAFNDVRVATYQVTRGLASLNSCVDPIFLYLAGDTFR 333
Qy 387 STAVLTXXNDLREOGQPSQ 406
Db 334 RLSRATRKASRRSEANLQSK 353
-----
RESULT 24
P2YR_BOVIN STANDARD; PRT; 373 AA.
AC P48042;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purinergic receptor).
GN P2RY1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic endothelium;
RX MEDLINE=95352058; PubMed=7626079;
RA Henderson D.J., Elliot D.G., Smith G.M., Webb T.E., Dainty I.A.;
RT "Cloning and characterisation of a bovine P2Y receptor.";
RL Biochem. Biophys. Res. Commun. 212:648-656(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Corpus callosum;
RX MEDLINE=99064562; PubMed=9848096;
RA Deng G., Matute C., Kumar C.K., Fogarty D.J., Miledi R.;
RT "Cloning and expression of a P2Y purinoceptor from the adult bovine
RL corpus callosum.";
RC Neurobiol. Dis. 5:259-270(1998).
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87628; CAA60958.1; -
DR EMBL; U34041; AAC78275.1; -
DR HSSP; P34996; 1DDO.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECP FL 1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 74
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 127 147
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 188
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 219 238
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 266 285
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 328
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 202 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42287 MW; 9270A7175C0BDA76 CRC64;
```

Query Match 18.5%; Score 398; DB 1; Length 373;
Best Local Similarity 28.5%; Pred. No. 2e-15;
Matches 111; Conservative 66; Mismatches 136; Indels 76; Gaps 15;

QY 45 DTNNLAKPTLPIKTRGAPPNSPEEPFSALEGWTGATI--TVKIKCPESASGHLHVKN 102
DB 14 DTAFLADP-----GSP-----WGNSTVTSTAASPFKCA-----43

QY 103 TMGVLTSLSLTKLIPAIYLLVVGVPANAVLWMLFFTRSG-ITVFYTNLAIDFL 160
DB 44 ----LTKTGFQYVLPVAVLYVFTIGLGNVAIMVFHMKPWSGISVYMFNLALADFL 99

QY 161 FCVTLPEKIAHLGNWVGEVLCRATTVFYGNMYCSILLACISINRYLAIVHPFTY 220
DB 100 YVLTLPALIFYFNKTDWIGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVPLKS 159

QY 221 RGLPKHTYALVTCGLWATVFLMPLFFILKQEY----LVQPDITTCCHDVHNTCESS 276
DB 160 LGRLEKKNVAVISLVMLVVGISPIL-----FYSGTGIRKNKTIICYDT-----TSDE 209

QY 277 FQLYYFI-----SLAFFGELIPFVLLIYCYAAIIRTNAYD-----HRLWVYKASLL 324
DB 210 YLSRYFYSMCTTIVAM--FCVPLVLILGCGYGLIVRALIYKDLDNSPLRRKSIYLV--II 264

QY 325 ILVFTICFAPSNIILIH-HANYVYNN-----TDGLYFYIYALCGLSGLNSCLDPELY 377
DB 265 VLTVFVAVSYIPFHWKTNLRLDQTPWCAFNDRVYATYQVTRGLASLNSCVDPILY 324

QY 378 FLMSKTRNHSTAYLTXXXNDLREQQPSQ 406
DB 325 FLAGDTFRRLSRATRKASRRSEANLQSK 353

RESULT 25
P2YR MOUSE
ID P2YR MOUSE STANDARD; PRT; 373 AA.

AC P49650;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 1 (ATP receptor) (P2Y1) (Purine receptor).
GN P2RY1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX TISSUE=Insulinoma;
RX MEDLINE=95298025; PubMed=7779087;
RA Tokuyama Y., Hara M., Jones E.M.C., Fan Z., Bell G.I.;
RT "Cloning of rat and mouse P2Y purinoceptors."
RL Biochem. Biophys. Res. Commun. 211:211-218(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Leon C.;
RT "Thromborectance in P2Y1 receptor knockout mice."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR FOR EXTRACELLULAR ADENINE NUCLEOTIDES SUCH AS
CC ATP AND ADP. IN PANCREATIC ISLETS, MAY MEDIATE SOME OF THE EFFECTS
CC OF EXTRACELLULAR ATP ON INSULIN SECRETION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U22829; AAA91302.1; .
DR EMBL; AJ245636; CAB57317.1; .
DR HSSP; P34996; 1DDO.
DR MGD; MGI:105049; P2ry1.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1.1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
DB DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 53 74 1 (POTENTIAL).
FT DOMAIN 75 87 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 88 109 2 (POTENTIAL).
FT DOMAIN 110 126 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 127 147 3 (POTENTIAL).
FT DOMAIN 148 166 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 167 188 4 (POTENTIAL).
FT DOMAIN 189 218 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 219 238 5 (POTENTIAL).
FT DOMAIN 239 265 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 266 285 6 (POTENTIAL).
FT DOMAIN 286 303 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 304 328 7 (POTENTIAL).
FT DOMAIN 329 373 CYTOPLASMIC (POTENTIAL).
FT DISULFID 124 202 BY SIMILARITY.
FT CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 113 113 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 373 AA; 42212 MW; 944125B9F4560BB3 CRC64;

Query Match 18.4%; Score 397.5; DB 1; Length 373;
Best Local Similarity 31.1%; Pred. No. 2.2e-15;
Matches 98; Conservative 61; Mismatches 117; Indels 39; Gaps 11;

QY 116 IPAYILLVVGVPANAVLWMLFFTRSG-ITVFYTNLAIDFLFCVTLPEKIAVHLN 174
DB 54 LPAYILLVFIIGLGNVAIMVFHMKPWSGISVYMFNLADFLVLTLPALIFYFN 113

QY 175 GNNVFGVLCRATTVFYGNMYCSILLACISINRYLAIVHPFTVGLPKHYVALVTCG 234
DB 114 KTDWIFGDAMCKLQRFIFHVNLYGSILFLTCISAHRYSGVVYPLKSLGRKKNAIYVS 173

QY 235 LWATVFLMPLFFILKQEY----LVQPDITTCCHDVHNTCESSSPQLYFI-----SL 285
DB 174 LWLIVVVAISPIL-----FYSGTGIRKNKTIICYDT-----TSNDYLSRYFYSMCTTV 223

QY 286 AFEGFLIPFVLLIYCYAAIIRTL--NAYDH-----RMLWYVYKASLLIIVFTICFAPSNI 338

DB 224 AM--FCIPLVLIIGCGYGLIVKALIYNDLNSPLRRKSIYLV--IIVLTVFVAVSYIPFH 278

QY 339 ILIIH-HANYVYNN-----TDGLYFYIYALCGLSGLNSCLDPELYFLMSKTRNHSTAYL 391
DB 279 MKTNLRLDQTPWCAFNDRVYATYQVTRGLASLNSCVDPILYFLAGDTFRRLSR 338

QY 392 TKXXNDLREQQPSQ 406
DB 339 TRKASRRSEANLQSK 353

RESULT 26
P2Y9 HUMAN
ID P2Y9 HUMAN STANDARD; PRT; 370 AA.

AC Q99677; O15132;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 9 (P2Y9) (Purine receptor 9) (G protein-coupled
GN receptor GPR23) (P2Y5-like receptor).
GN GPR23 OR P2RY9.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```
Db 123 ATAHFQMLHITSVRSAWILCGIIV--VFIMASSGLLLKHGOE---KQNTTLQFELN-- 175
Qy 271 CESSSPFOLYYFISLAFGLPFLPVLIIYCYAAAIIRTNAYDRWLWVVKASL----- 323
Db 176 LQKFNVLNLYALG-VGFLLPFLITCYLLIIRVL-----LKEIPESGPRD 224
Qy 324 -----LILVIFTICFAPSNIILIIHHANYNN-TDGLYFIYLIACLCGLNSCL 372
Db 225 AQRKALTTIVAMIIIFLLCFPLHYALRTIHLVTWDAQSCMDLKHATVITLTLAANSCF 284
Qy 373 DPFLYF 378
Db 285 NPFLY 290

RESULT 29
CLT2_PIG
ID CLT2_PIG STANDARD; PRT; 345 AA.
AC Q9SN03;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyln leukotriene receptor 2 (CysLTR2).
GN CysLTR2 OR CysLTR2.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kanohara M., Takasaki J., Matsumoto M., Matsumoto S., Saito T.,
RA Ohishi T., Soga T., Matsushima H., Furuichi K.,
RT "Characterization of the cloned rat and porcine cysteinyl leukotriene
RT receptors";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes. The response is
CC mediated via a G-protein that activates a phosphatidylinositol-
CC calcium second messenger system (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB052662; BAB60817.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 43
FT TRANSMEM 44 64
FT DOMAIN 65 73
FT TRANSMEM 74 94
FT DOMAIN 95 124
FT TRANSMEM 125 145
FT DOMAIN 146 154
FT TRANSMEM 155 175
FT DOMAIN 176 205
FT TRANSMEM 206 226
FT DOMAIN 227 246
FT TRANSMEM 247 267
FT DOMAIN 268 287
FT TRANSMEM 288 308
FT DOMAIN 309 345
FT DISULFID 112 188
FT CARBOHYD 20 20
FT CARBOHYD 29 29
```

```
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 345 AA; 39410 MW; 5D1B1FB9BB95905 CRC64;

Query Match 17.68; Score 379.5; DB 1; Length 345;
Best Local Similarity 28.74; Pred. No. 1.9e-14;
Matches 97; Conservative 68; Mismatches 136; Indels 47; Gaps 13;

Qy 92 ESASHLHVKNATMGYLTSSLTSLIPAIYLLVFGVGVGPNNAVTLMLPFR--TRSICTTV 149
Db 18 EPNSTLGNHNSRSCITTENKREFYPIVLIIVFINGALGNQFSIY-VLPKPKKSTSVV 76
Qy 150 FYTNLAIADELFQVTLFPFKIAYHLNGNNWVGEVLCRATTVIFGNNMYCSILLACISIN 209
Db 77 FMLNLAIISDLLFTITLPPRVYVLRGSNXIFGDTPCRIMSYMVNNYSSIIYFLTVLSV 136
Qy 210 RYLAIVHPFTYRGLPKHITYALVTCGLVWATVFLYMLPFFILKQBYLV-----QPDITT 263
Db 137 RFLATVHPFRLHTTTSIKNAWILCGVIW-----IFIMASVTLLKNGSEQKDNVTL 187
Qy 264 CHDVHNTCESSPFQLYYFISLAFGLPFLIIPVLIYCYAAAIIRTNAYD-----HR 314
Db 188 CLELNS--NKVTYKLTWNYVALV-VGFVLPFGTILSICYLIIIRALLKVEVPESEGLRSHR 244
Qy 315 WLMYVVKASLLI---LVIFTICFAPSNIILIIHHANYNN-TDGLYFIYLIACLCGLNS 370
Db 245 -----KALITVITIAIIIFLLCFPLHYVLRTRLHLEWKADCKDRLHKAVALTLAAANS 299
Qy 371 CLDPFLYFLMSKTRNHSHTAYLTXXXNDLREQQPSQRT 408
Db 300 CFNPFYLYFAGEN-----FKDRLKSALR-KGRP-QKT 329

RESULT 30
P2Y4_RAT
ID P2Y4_RAT STANDARD; PRT; 361 AA.
AC O35811;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 4 (P2Y4).
GN P2Y4 OR P2Y4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Liver;
RA Bogdanov Y.D., Wildman S., King B.F., Burnstock G.;
RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Brain;
RA MEDLINE=98421785; PubMed=9751165;
RA Webb T.E., Henderson D., Roberts J.A., Barnard E.A.;
RT "Molecular cloning and characterization of the rat P2Y4 receptor.";
J. Neurochem. 71:1424-1434(1998).
CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
CC activate a phosphatidylinositol-calcium second messenger system.
CC Not activated by ADP or UDP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels. In brain,
CC higher expression in the pineal gland and ventricular system.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```



```

FT DOMAIN 64 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 93 2 (POTENTIAL).
FT DOMAIN 94 123 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 124 144 3 (POTENTIAL).
FT DOMAIN 145 153 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 154 174 4 (POTENTIAL).
FT DOMAIN 175 204 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 205 225 5 (POTENTIAL).
FT DOMAIN 226 245 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 246 266 6 (POTENTIAL).
FT DOMAIN 267 286 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 287 307 7 (POTENTIAL).
FT DOMAIN 308 346 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 187 BY SIMILARITY.
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 181 181 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 346 AA; 39635 MW; EB544A42DDC5EE4 CRC64;

Query Match 17.5%; Score 376.5; DB 1; Length 346;
Best Local Similarity 30.1%; Pred. No. 2.8e-14;
Matches 93; Conservative 62; Mismatches 129; Indels 25; Gaps 9;

Qy 82 TITVKKCPESASHLVKNATMGVLTSSLSKLPALYLLVFGVPANAVTLMLFFR 141
Db 11 SISVSEMEPNCTFNNSNRNCTI---ENFRREFPVIYLLIFFGVGLNGLSIY-VFLQ 65

Qy 142 --TSICITVTVYTLATADFLCVTLPEKIAHYLNGNNWGEVLCRATTIVIFGNMYS 199
Db 66 PYKSTSVNFMULNLSDLFIPTLPFRADYLRGSNWIFEDLACRIMSLSLVNMYSS 125

Qy 200 ILLACISINRYLAIVHPFVYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQP 259
Db 126 IYFVLTVSVRFLAVHPFRLHVTISRAMLGLIWN---ILLWASSIMLLDSSGSGQ 182

Qy 260 DITCHDVHNTCESSPQLYFISLAFGLIPFLVLIYCYAAILRTLNAYD----- 312
Db 183 SVTSCLELN--LYKIAKLQTNWYIALV-VGCLLPFFTLISICYLLIIRVLLKVEVPESGLR 239

Qy 313 --HRLWVYKASLLILVFTICFAPSNIILIIHANYYYN-NTDGLFYIYLIACLSLN 369
Db 240 VSHRKA--LTIIITLIIFLCFLPYHTLRTVHLTTWKVGLCKDLRHLKALVITLALAA 297

Qy 370 SCLDPFLYF 378
Db 298 ACPNPLLYY 306

```

RESULT 32

```

P2Y5 HUMAN
ID P2Y5_HUMAN STANDARD; PRT; 344 AA.
AC P43657; O15133;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P2Y purinoceptor 5 (p2Y5) (Purinergic receptor 5) (RB intron encoded
DE G-protein coupled receptor).
GN P2Y5
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bohm S.K., Trummel A., Khitin L.M., Kong W., Payan D.G., Bunnett N.W.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RX [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=94063891; PubMed=7902321;
RA Tochida J., McGee T.L., Paterson J.C., Eagle J.R., Tucker S.,
RA Yandell D.W., Dryja T.P.;
RT "Complete genomic sequence of the human retinoblastoma susceptibility

```

```

gene.";
Genomics 17:535-543 (1993).
[3]
IDENTIFICATION.
Church G.M.;
Unpublished observations (OCT-1994).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -!- CAUTION: REF.2 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 31.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; L11910; -; NOT ANNOTATED_CDS.
CC EMBL; AF000546; AAB62190.1; -.
CC HSP; P34996; 1DDD.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC Pfam: PF00001; 7tm1; 1.
CC PROSITE; PS00237; G-PROTEIN RECP FL 1; 1.
CC PROSITE; PS00262; G-PROTEIN RECP FL 2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
CC Palmitate.
DOMAIN 1 19 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 20 46 1 (POTENTIAL).
FT DOMAIN 47 55 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 56 79 2 (POTENTIAL).
FT DOMAIN 80 92 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 93 112 3 (POTENTIAL).
FT DOMAIN 113 133 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 134 154 4 (POTENTIAL).
FT DOMAIN 155 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 209 5 (POTENTIAL).
FT DOMAIN 210 227 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 228 253 6 (POTENTIAL).
FT DOMAIN 254 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 292 7 (POTENTIAL).
FT DOMAIN 293 344 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 284 284 PALMITATE (BY SIMILARITY).
FT DISULFID 89 168 POTENTIAL.
FT CONFLICT 33 33 V -> I (IN REF. 2).
SQ SEQUENCE 344 AA; 39377 MW; 7E31155AFFB499AB CRC64;

Query Match 17.4%; Score 376; DB 1; Length 344;
Best Local Similarity 30.9%; Pred. No. 3e-14;
Matches 96; Conservative 55; Mismatches 122; Indels 38; Gaps 10;

Qy 101 NATGWYLTSSLSKLPALYLLVFGVPANAVTLMLPFRTRISIC-----TTVPYTN 153
Db 5 NSSGCFYNDSEFKYTLGCMESVFLGLVSNCAIYIF-----ICVLKVRNRTTYMIN 58

Qy 154 LATADFLCVTLPEKIAHYLNGNNWGEVLCRATTIVIFGNMYSCTILLACISINRYLA 213
Db 59 LAMSDLLFVFLTPFRI-FYETTRNPPGDLCKISVNLFTYNNMYSILFTICISVDRLA 117

Qy 214 IVHPFVYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQBYLVQPDIITCHDVHNTCES 273
Db 118 IVYFESKTLRTKRNKAVICTGWLTVIGGSAP-----AVFQSTHSGQNNASEACFE 170

Qy 274 SSFPQLY--YFISLAF-----GFLIPFLVLIYCYAAILRTLN---AYDHRWLWYVKASLL 324
Db 171 NFPEATWYLSRIVIFIEIVGFFIPLILNVTCSSMWLKTLPVTLRSKINKTKVLM 230

Qy 325 I---LVITFCFAPSNIILIIH---HANYYYNT--DGLFYIYLIACLSLNSCLDPFL 376
Db 231 IFVHLIIFCFYFYNILYLSVTRTQTFFVNCVSAVARTMYPTILCIAVSNCCFDPV 290

```

```

QY 377 YFLMSKTRNHS 387
Db 291 YFTSDTIQNS 301

RESULT 33
CLT1_PIG STANDARD; PRT; 340 AA.
AC Q95N02; 2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyll leukotriene receptor 1 (CysLTR1).
GN CysLTR1 OR CysLTR1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Takasaki J., Kanohara M., Saito T., Matsumoto M., Matsumoto S.,
RA Ohishi T., Soga T., Matsushime H., Furuichi K.;
RT "Characterization of cloned rat and porcine cysteinyl leukotriene
RT receptors.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC constriction of the microvascular smooth muscle during an
CC inflammatory response. This response is mediated via a G-protein
CC that activates a phosphatidylinositol-calcium second messenger
CC system (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB052686; BAB60826.1; -
CC InterPro; IPR000276; GPCR_Rhodopsin.
CC Pfam; PF00001; 7tm.1.1.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR_F1_1; FALSE_NEG.
CC PROSITE; PS0262; G-PROTEIN RECEPTOR_F2_1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein.
KW DOMAIN 1 31 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 32 52 1 (POTENTIAL).
FT DOMAIN 53 60 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 61 81 2 (POTENTIAL).
FT DOMAIN 82 109 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 110 130 3 (POTENTIAL).
FT DOMAIN 131 144 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 145 165 4 (POTENTIAL).
FT DOMAIN 166 196 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 197 217 5 (POTENTIAL).
FT DOMAIN 218 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 254 6 (POTENTIAL).
FT DOMAIN 255 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 300 7 (POTENTIAL).
FT DOMAIN 301 340 CYTOPLASMIC (POTENTIAL).
FT DISULFID 99 176 BY SIMILARITY.
FT CARBOHYD 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 340 AA; 38986 MW; 54F9372A121CE413 CRC64;

Query Match 17.4%; Score 375.5; DB 1; Length 340;
Best Local Similarity 31.5%; Pred. No. 3.2e-14;
Matches 106; Conservative 56; Mismatches 136; Indels 37; Gaps 12;

QY 99 VKNATMGYLTS- - - - - LSKLIPAIYLYLVFVGVANAVTLMLF- - - - - FRTSICIT 148
Db 4 VRNLTVSCASNTCDTIDDFRNQVSTLYSMITVVVFGNGFVLVLKTYHEKS-AYQ 62
QY 149 VFYTNLAIAADFLFCVTLPLPKIAYHLNGNNWVGEVLCRATTVIFYGNMVCISILLACISI 208
Db 63 VYMINLAVADLLCVCVTLPLRVVYVHKGIWLGDFLCRLSTVALYVNLVYCSIFFMTAMSF 122
QY 209 NRYLAIVHPFTYRGLPKHTYALVTCGLVWATVYMLPFFILKQEVYLVOPDITTCDDVH 268
Db 123 FRCIAIVFPVQINILTHKKAKIVCIAIWFILVLTSSPF- - - - - LMSTYKOEKNNTKCFEPP 180
QY 269 NTCESSPPFOLYYFISLAFEGFLIPFVLIYCVAAIIRTL- - - - - NAYDHRMLWYVKASILL 326
Db 181 QXNQAKYHVLVHLVSL- - - - - FVGFIIPFVLIYCVYMIILLKNSMKNISRRKKAICMII 239
QY 327 VI- - - - - FTICFAPSNIILIIHHANYNNNT- - - - - DGLYFI- - - - - YLIAICLSGLNSCLDPFLY 377
Db 240 VVTAAPFLISFMPYHIQRTI- - - - - HLHFLHNDTKGCDVLMQKSVXITLSLAASNCDFDPLLY 298
QY 378 FL- - - - - MSKTRNHSTA- - - - - YLTXXNDLREQG 403
Db 299 FFSGGNFRGLSTFRKHSLSUTMTYVPKKTSLEPKAQ 335

RESULT 34
CLT1_HUMAN STANDARD; PRT; 337 AA.
AC Q9Y271;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cysteinyll leukotriene receptor 1 (CysLTR1) (Cysteinyll leukotriene D4
DE receptor) (LTD4 receptor) (HG55) (HMTMF81).
GN CysLTR1 OR CysLTR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tonsil;
RX MEDLINE=99318129; PubMed=10391245;
RA Lynch K.R., O'Neill G.P., Liu Q., Im D.-S., Sawyer N., Metters K.M.,
RA Coulombe N., Abramovitz M., Figueroa D.J., Zeng Z., Connolly B.M.,
RA Bai C., Austin C.P., Chateaufneuf A., Stocco R., Greig G.M.,
RA Kargman S., Hooks S.B., Hosfield E., Williams D.B. Jr.,
RA Ford-Hutchinson A.W., Caskey C.T., Evans J.P.;
RT "Characterization of the human cysteinyl leukotriene CysLTR1
RT receptor.";
RL Nature 399:789-793 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood monocytes, Leukocyte, and Spleen;
RX MEDLINE=99393629; PubMed=10462554;
RA Sarau H.M., Ames R.S., Chambers J., Ellis C., Elshourbagy N.,
RA Foley J.J., Schmidt D.B., Muccitelli R.M., Jenkins O., Murdoch P.R.,
RA Herriy N.C., Halsey W., Sathe G., Muir A.I., Nuthulaganti P.,
RA Dytko G.M., Buckley P.T., Wilson S., Bergsma D.J., Hay D.W.P.;
RT "Identification, molecular cloning, expression, and characterization
RT of a cysteinyl leukotriene receptor.";
RL Mol. Pharmacol. 56:657-663 (1999).
RN [3]
RP -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
CC bronchoconstriction of individuals with and without asthma.
CC Stimulation by LTD4 results in the contraction and proliferation
CC of smooth muscle, edema, eosinophil migration and damage to the
CC mucus layer in the lung. This response is mediated via a G-protein
CC that activates a phosphatidylinositol-calcium second messenger
CC system. The rank order of affinities for the leukotrienes is LTD4
CC >> LTE4 = LTC4 >> LTBA.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC spleen and peripheral blood leukocytes. Lower expression in

```


CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF329272; AAK16715.1; -
DR EMBL; AF329272; AAK16716.1; -
DR EMBL; AF205830; AAK15433.1; -
DR EMBL; AF363370; AAK73047.1; -
DR EMBL; AB044087; BAA96809.1; -
DR HSP; P34996; 1DDO.
DR MGD; MGI:1926218; Cyslrl.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; FALSE_NEG.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Alternative splicing.
FT DOMAIN 1 43 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 44 64 1 (POTENTIAL).
FT DOMAIN 65 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 93 2 (POTENTIAL).
FT DOMAIN 94 121 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 122 142 3 (POTENTIAL).
FT DOMAIN 143 156 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 157 177 4 (POTENTIAL).
FT DOMAIN 178 208 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 209 229 5 (POTENTIAL).
FT DOMAIN 230 245 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 246 266 6 (POTENTIAL).
FT DOMAIN 267 291 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 292 312 7 (POTENTIAL).
FT DOMAIN 313 352 CYTOPLASMIC (POTENTIAL).
FT DISULFID 111 188 BY SIMILARITY.
FT CARBOHYD 15 15 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 26 26 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 13 MISSING (IN ISOFORM 2).
FT CONFLICT 176 176 Y -> D (IN REF. 3).
SQ SEQUENCE 352 AA; 40715 MW; 58DC94B3F1CD0CAB CRC64;

Query Match 17.2%; Score 371.5; DB 1; Length 352;
Best Local Similarity 30.5%; Pred. No. 5.3e-14;
Matches 107; Conservative 49; Mismatches 120; Indels 75; Gaps 12;
QY 108 TSSLSTKLI-----PAIYLVFVGVGPANAVTLWMLF--FRTRISCTTV 149
DB 17 TENLTSLINNTCHTDIDFRNQVSTMYSVISVGVFGFNSFVLYVLIKTVHEKS-AFQV 75
QY 150 FYTNLAIDFLFCVTLTPKIAHYHLNGNNWVEVLCRATTIFYGNMYCSILLACISIN 209
DB 76 YMIHLAIDFLFCVTLPLRVVYVYHKGWLFGLRLTLYVALYVNYLCSIFFMTAMSF 135
QY 210 RYLAIVHPFTYRGLPKHYALVTCGLVWATVFLKQVYLVQPDITTCVDVHN 269
DB 136 RCVAIVFPQVINLVTKQKARFVIGIWFILTSFPLMYKS--YQDKNTKCFPPQ 193
QY 270 TCESSSPQLYFISLAFPGELIPFLIYCAIIRTL-----NAYDHRWLWYKAS 322
DB 194 NNQAKYVILHLYVSL-FFGIIPFTIIVCYTWIILTLKNTMKKNPSRR-----KAI 247
QY 323 LLILVI---FPICAPSNIIILHHYNNYNTDGLFYI-----LIAL 363
DB 248 GMIIIVTAAFLVSPMPYHIQRTIH-----LHLHSETRPCDSVLRMQKSVVITL 296
QY 364 CLGSNCLDPLFYFL-----MSKTRNHSTA---YLTXXNDLREOQ 403
DB 297 SLAASNCDFLLIFFSGCNFRRLSTFRKHSLSMTYVPPKKASLPEKGE 347

RESULT 36
P2Y4_MOUSE

ID P2Y4_MOUSE STANDARD; PRT; 361 AA.
AC OSJJS7;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 4 (P2Y4).
GN P2RY4 OR P2Y4R.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ;
RX MEDLINE=21185993; PubMed=11290369;
RA Suarez-Huerta N.; Pouillon V.; Boeynaens J.-M.; Robaye B.;
RT "Molecular cloning and characterization of the mouse P2Y4 nucleotide
receptor.";
RL Eur. J. Pharmacol. 416:197-202(2001).
CC -1- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
activate a phosphatidylinositol-calcium second messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in the liver, intestine, stomach,
bladder and lung.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ277752; CAB91043.1; -
CC HSP; P34996; 1DDO.
DR MGD; MGI:1926594; P2RY4.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 30 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 31 58 1 (POTENTIAL).
FT DOMAIN 59 68 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 69 91 2 (POTENTIAL).
FT DOMAIN 92 108 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 109 127 3 (POTENTIAL).
FT DOMAIN 128 149 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 150 170 4 (POTENTIAL).
FT DOMAIN 171 192 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 193 218 5 (POTENTIAL).
FT DOMAIN 219 242 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 243 265 6 (POTENTIAL).
FT DOMAIN 266 283 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 284 305 7 (POTENTIAL).
FT DOMAIN 306 361 CYTOPLASMIC (POTENTIAL).
FT DISULFID 104 181 BY SIMILARITY.
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 361 AA; 41034 MW; 3E8EA84B65BC0A20 CRC64;

Query Match 17.2%; Score 370; DB 1; Length 361;
Best Local Similarity 31.4%; Pred. No. 6.6e-14;
Matches 98; Conservative 49; Mismatches 129; Indels 36; Gaps 11;
QY 115 LIPAIYLVFVGVGPANAVTLWMLF--FRTRISCTTVFT--NLAIADFLFCVTLTPKIAHYHL 173
DB 33 LLPLSYAVFVGLGALNAPTLLFLRLRPWDATATYMFHLASDTLYLSLPLTVYYA 92
QY 174 NGNNWVEVLCRATTIFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHYALVTC 233
DB 93 ARNHWPFGTGFCFKFVRFVFLFYWNLYCSVLFTLCISVHYRMYGICHPLRAIRWGRPRFAGLLC 152

QY 234 GLVWATVFLYMLP--FFILKQEYLVQPDIT--CHDVHNTCESSPPQLYY--FISLAF 287
 DB 153 LGVWLVAGCLVPLFFV-----TTNANGTILCHD--TTLPEFDHYVFSSTIMVLL 204
 QY 288 FGLIPFLVLLIYCYAAIIRTL-----NAYDHRMLWYVKASLLILVITFCFAPSNIILI 341
 DB 205 FGF--PFLITLVCGLMARRLYRPLPGAGQSSRLSLRTIAVLTTFVAVCFVPHITRT 262
 QY 342 IHHANYNNYNN---TDGLYFYTLIALCLGSLNSCLDPFLY--FLMSKTRNHSTAYLTKXXN 396
 DB 263 IYLLARULNACRVLNIVNVYKVRPLASANSCLDPVLYLFTGDKYRQ-----LQ 314
 QY 397 DLREQQPSQRT 408
 DB 315 QLCRGSTPKRRT 326

RESULT 37

P2Y2_HUMAN STANDARD; PRT; 377 AA.
 AC P41231; Q96EM8;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
 DE (Purinoceptor receptor)
 GN P2RY2 OR P2RU1
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 .X NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 RP TISSUE=Airway epithelium;
 RC MEDLINE=94211846; PubMed=8159738;
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
 RT "Cloning and expression of a human P2U nucleotide receptor, a target
 RT for cystic fibrosis pharmacotherapy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3275-3279 (1994).
 [2]
 RP REVISIONS.
 RX MEDLINE=95108098; PubMed=7809171;
 RA Parr C.E., Sullivan D.M., Paradiso A.M., Lazarowski E.R., Burch L.H.,
 RA Olsen J.C., Erb L., Weisman G.A., Boucher R.C., Turner J.T.;
 RT "Cloning and expression of a human P2U nucleotide receptor, a target
 RT for cystic fibrosis pharmacotherapy.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:13067-13067 (1994).
 [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney, and Leukocyte;
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
 CC activate a phosphatidylinositol-calcium second messenger system.
 CC The affinity range is UTP = ATP > ATP-gamma-S > 2-methylthio-ATP
 CC = ADP.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: SPLEEN, KIDNEY, LIVER, LUNG, HEART AND
 CC BRAIN.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U07225; AAC04923.1; -;
 CC EMBL; BC012104; AAH12104.1; -;

DR EMBL; BC028135; AAH28135.1; -;
 DR PIR; A54946; A54946.
 DR HSP; P34996; IIDD.
 DR Genew; HGNC:8541; P2RY2.
 DR MM; 600041;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PRO0237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
 DR PROSITE; PS00262; G-PROTEIN RECEPT_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein.
 FT DOMAIN 1 32 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 33 59 1 (POTENTIAL).
 FT DOMAIN 60 70 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 71 93 2 (POTENTIAL).
 FT DOMAIN 94 110 3 (POTENTIAL).
 FT TRANSMEM 111 129 3 (POTENTIAL).
 FT DOMAIN 130 152 4 (POTENTIAL).
 FT TRANSMEM 153 172 4 (POTENTIAL).
 FT DOMAIN 173 194 5 (POTENTIAL).
 FT TRANSMEM 195 220 5 (POTENTIAL).
 FT DOMAIN 221 246 6 (POTENTIAL).
 FT TRANSMEM 247 269 6 (POTENTIAL).
 FT DOMAIN 270 287 7 (POTENTIAL).
 FT TRANSMEM 288 309 7 (POTENTIAL).
 FT DOMAIN 310 377 7 (POTENTIAL).
 FT CARBOHYD 9 9 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 13 13 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 106 183 BY SIMILARITY.
 FT CONFLICT 312 350 R -> S (IN REF. 3; AAH12104).
 FT CONFLICT 350 350 E -> G (IN REF. 1).
 FT CONFLICT 359 359 S -> F (IN REF. 1).
 SQ SEQUENCE 377 AA; 42289 MW; EE557A857A269AC6 CRC64;
 Query Match 17.0%; Score 366; DB 1; Length 377;
 Best Local Similarity 29.2%; Pred. No. 1.1e-13;
 Matches 91; Conservative 52; Mismatches 135; Indels 34; Gaps 9;
 QY 115 LIPAIYLLVFGVGPANAVTLWMLFFRTS-ICTTVFYTNLAIDFLFCVTLPLPKIAYHL 173
 DB 35 LLPVSTGVGVLCGLNAVALYIFLCRLKTFWNASTTYMFHLAVSDALYAAASLLPLLVYYA 94
 QY 174 NGNNVFGVGLCRATTVIFYGNMYCSILLIACISINRYLAIVHPFTYRGLPKHTYALVTC 233
 DB 95 RGDHWFSTVLCVLRFLFYTNLYCSILFICISVHRCGLVRLSLRWGRARYARRVA 154
 QY 234 GLVWATVFLYMLP--FFILKQEYLVQPDIT--CHDVHNTCESSPPQLYYFISLAF 291
 DB 155 GAVWLVLAGQAPVLYFVTT-----ARGGRVTCHDT-SAPELFSRFVAYSSVMGLGLL-FA 208
 QY 292 IPEVLIYCYAAIIR-----TLNAYDHRMLWYVKASLLILVITFCFAPSNIILI 343
 DB 209 VPFAVLVYVLMARLLKPAVGTSGCLPRACKSVRTTAVLAVFALCPLPHVTRTY 268
 QY 344 HANYNNYNTD-----GLYFYIYIALCLGSLNSCLDPFLYFLMSKTRNHSTAYLTKXXN 396
 DB 269 ---YSRSLDLSCHTLNAINMAYKVRPLASANSCLDPVLYFLAGQ-----LVRFR 318
 QY 397 DLREQQPSQRT 408
 DB 319 DAKPPTGSPAT 330
 RESULT 38
 ID CLT1_RAT STANDARD; PRT; 339 AA.
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cysteinyl leukotriene receptor 1 (CysLTR1).
 GN CysLTR1 OR CysLTR1.
 OS Rattus norvegicus (Rat).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
[1]
SEQUENCE FROM N.A.
Takasaki J., Kamohara M., Saito T., Matsumoto M., Matsumoto S.,
Ohishi T., Soga T., Matsushime H., Furuichi K.;
"Characterization of cloned rat and porcine cysteinyl leukotriene
receptors";
Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for cysteinyl leukotrienes mediating
contraction of the microvascular smooth muscle during an
inflammatory response. This response is mediated via a G-protein
that activates a phosphatidylinositol-calcium second messenger
system (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; AB052685; BAB60825.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G-PROTEIN RECP F1.1; FALSE_NEG.
DR PROSITE; PS0262; G-PROTEIN RECP F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 30
FT TRANSMEM 31 51
FT DOMAIN 52 59
FT TRANSMEM 60 80
FT DOMAIN 81 108
FT TRANSMEM 109 129
FT DOMAIN 130 143
FT TRANSMEM 144 164
FT DOMAIN 165 195
FT TRANSMEM 196 216
FT DOMAIN 217 232
FT TRANSMEM 233 253
FT DOMAIN 254 278
FT TRANSMEM 279 299
FT DOMAIN 300 339
FT TRANSMEM 340 359
FT DISULFID 98 175
FT CARBOHYD 6 6
FT CARBOHYD 171 171
SQ SEQUENCE 339 AA; 39143 MW; 281B41DF050DF8A CRC64;
Query Match 16.9%; Score 365.5; DB 1; Length 339;
Best Local Similarity 31.5%; Pred. No. 1.1e-13;
Matches 102; Conservative 46; Mismatches 117; Indels 59; Gaps 11;
QY 119 IYLVFVGVDPANAVTLMLF--PRTSICITVYTNLAIDFLCVTLFPKIAHLNGN 176
DB 31 MYSIMSVGVFGNSFVLYVLIKTTHKES-AFQVYMINLAIDLCLVCTPLRVVYVYHK 89
QY 177 NWVGEVLCAITTVIFVGNMYSILLIACISINYLAVHPFTYVRGLPKHTYALVTCGLV 236
DB 90 KWFQDFLCRLTTVALYVNLVCSIFFMTAMSFRCVAVLFPVQINLVTKKARFVCVGI 149
QY 237 WATVFLYMLPFIILKQEVYLVQPDITTDCHDVHNTCESSSPQLYYFISLAFFGLIPFVL 296
DB 150 WIFVILTSSPELLSKS--YQDEKNTKCFEPQDKQTKYVLYVLYVSL-IFGFIIPFT 206
QY 297 ILYCYAAIIRL-----NAYDRWLWYVKASLLILVI---FTICAPSNIILIIHAN 346
DB 207 IIVCYTMIIILLKNTMKNLPSRR-----KAIGMIIVVTAFLVSPMPYIQRAIH--- 258
QY 347 YYNNNTDGLYFIY-----LIALGLGSLNSCLDPLFLVFL-----MSK 382

Db 259 -----LHFLHSETRSCDSVLRMQKSVVITLSLAASCCCEPFLYFFSGGNFRRRLST 310
QY 383 TRNHSTA---YLTXXNDLREQG 403
Db 311 FRKHSLSMTYTPKKASLPKGE 334
RESULT 39
P2Y2_RAT STANDARD; PRT; 374 AA.
AC P41232;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE P2Y purinoceptor 2 (P2Y2) (P2U purinoceptor 1) (P2U1) (ATP receptor)
DE (Purinegic receptor).
GN P2RY2 OR P2RU1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95110548; PubMed=7811468;
RA Rice W.R., Burton F.M., Fiedelvey D.T.;
RT "Cloning and expression of the alveolar type II cell P2u-purinegic
receptor.";
RL Am. J. Respir. Cell Mol. Biol. 12:27-32(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Pituitary;
RA Chen Z.P., Krull N., Xu S., Levy A., Lightman S.L.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RX MEDLINE=98099857; PubMed=9437211;
RA Seye C.I., Gadeau A.P., Darot D., Dupuch F., Alzieu P., Capron L.,
RA Desgranges C.;
RT "Overexpression of P2Y2 purinoceptor in intimal lesions of the rat
aorta.";
RL Arterioscler. Thromb. Vasc. Biol. 17:3602-3610(1997).
CC -!- FUNCTION: Receptor for ATP and UTP coupled to G-proteins that
activate a phosphatidylinositol-calcium second messenger system.
CC The affinity range is UTP > ATP > ATP-gamma-S >> 2-methylchlo-ATP
= ADP.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U09402; AAA61565.1; -
DR EMBL; L46865; AAB02099.1; -
DR EMBL; U56839; AAC00048.1; -
DR HSP; P34996; 1DDD
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_Rhodpsn.
DR PROSITE; PS00237; G-PROTEIN RECP F1.1; 1.
DR PROSITE; PS0262; G-PROTEIN RECP F1.2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 32
FT TRANSMEM 33 59
FT DOMAIN 60 70
FT TRANSMEM 71 93
FT DOMAIN 94 110
FT EXTRACELLULAR (POTENTIAL).
FT CYTOPLASMIC (POTENTIAL).
FT EXTRACELLULAR (POTENTIAL).

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 11:59:21 ; Search time 26.0319 Seconds
(without alignments)
1506.722 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMILQISXRLRDTQVIKM.....AYLTXXNDLREQQPSQRT 408
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	606.5	28.1	399	2 I48705	proteinase activat
2	599.5	27.8	420	2 I51667	thrombin receptor
3	589.5	27.3	397	2 S65518	proteinase-activat
4	523	24.2	425	2 A37912	thrombin receptor
5	505.5	23.4	432	2 A43448	thrombin receptor
6	501.5	23.2	427	2 S17148	alpha-thrombin rec
7	424	19.7	362	2 S33733	G protein-coupled
8	410.5	19.0	361	2 B45680	G protein-coupled
9	406.5	18.8	308	2 I50241	G protein-coupled
10	399	18.5	373	2 J4737	G protein-coupled
11	398	18.5	373	2 J4162	P2Y receptor - bov
12	384.5	17.8	370	2 JCS549	heptahelical P2Y5
13	382	17.7	365	2 S68679	G protein-coupled
14	376	17.4	344	2 T09508	intron 17 purinerg
15	363.5	16.9	355	2 J01231	interleukin-8 rece
16	361	16.7	373	2 A47556	ATP receptor P2u -
17	356.5	16.5	362	2 JN0694	angiotensin II rec
18	356	16.5	375	2 A54946	p-2U nucleotide re
19	354	16.4	350	2 A39445	interleukin-8 rece
20	353	16.4	356	2 S42096	interleukin-8 rece
21	352	16.3	328	2 I55450	G protein-coupled
22	351.5	16.3	360	2 A53611	interleukin-8 rece
23	351	16.3	365	2 S68208	G protein-coupled
24	347	16.1	361	2 JCS653	G protein-coupled
25	346	16.0	341	2 S63666	platelet activatin
26	345.5	16.0	359	2 A48921	interleukin-8 rece
27	345.5	16.0	362	2 B57641	G protein-coupled
28	345	16.0	359	2 S15403	angiotensin II rec
29	344	15.9	372	2 S26667	G protein-coupled

30 343 15.9 359 2 I51372 angiotensin II rec
31 342.5 15.9 342 2 A40191 platelet-activatin
32 342.5 15.9 358 2 A53752 interleukin-8 rece
33 342 15.9 374 2 S32785 G protein-coupled
34 342 15.9 380 2 A48227 kappa opioid recep
35 341.5 15.8 342 2 S13638 platelet-activatin
36 341.5 15.8 369 2 JCS068 G protein-coupled
37 341 15.8 359 2 A48857 angiotensin II rec
38 341 15.8 359 2 JCS134 angiotensin II rec
39 341 15.8 380 2 JC2434 kappa opioid recep
40 340 15.8 341 2 S43252 platelet-activatin
41 340 15.8 359 2 JH0621 angiotensin II rec
42 340 15.8 380 2 S36143 kappa opioid recep
43 339.5 15.7 354 2 I53033 G protein-coupled
44 339.5 15.7 362 2 A57641 G protein-coupled
45 .339 15.7 353 2 JC2492 G protein-coupled

ALIGNMENTS

RESULT 1

I48705
Proteinase activated receptor 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C;Accession: I48705
R;Nystedt, S.; Larsson, A.K.; Aberg, H.; Sundelin, J.
J. Biol. Chem. 270, 5950-5955, 1995
A;Title: The mouse proteinase-activated receptor-2 cDNA and gene. Molecular cloning and
A;Reference number: I48705; MUID:95197620; PMID:7890726
A;Accession: I48705
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-399 <RES>
A;Cross-references: EMBL:Z48043; NID:g663020; PIDN:CAA88097.1; PID:g663021
C;Superfamily: ATP receptor P2u

Query Match 28.1%; Score 606.5; DB 2; Length 399;
Best Local Similarity 35.4%; Pred. No. 2.9e-43;
Matches 137; Conservative 69; Mismatches 126; Indels 55; Gaps 12;

Qy 22 ALIFAAAGLLLLPTFCQGMEN--DTNNIAKPTLPKTRGAPPNSFEBFPFSALEGWT 79
Db 5 SLAWLLGGITLLAASVSCSRNTENLAPGRNNSKGRSLIGRLTQPP-----IT 51
Qy 80 GAITVYK--IKCPESASHLHVKNATMGVLTSSLYKLIPAIYLLVYVGVGPANAVTLWM 137
Db 52 GKGVPEPGFSIDFSSAS-----ILTGKLTTFLEPVVYIIVFVIGLPSNGMALWI 101
Qy 138 LFRTRTSICTTVFY-TNLAIADFLFCVTLPFKIAYHLNGNWNVGFVCLCRATTVIFYGNM 196
Db 102 FLRTKKKPAVIYMANLADLLSVIFWFLKLSYHLHGNNWVYGEALCKVLGFFVGNM 161
Qy 197 YCSILLACISINRYLAIVHPFTYRGLPKHTYALVTGCL---VWATVFLYMLPFFILKQE 253
Db 162 YCSILFWTCLSVQRYWVIVNPM---GHPRKK-ANIAVGVSALWLLFLFTVPLIYVMKQT 217
Qy 254 YVLVQPDITTCVDVHTNCESSSPFQLY-----YFISLAFGFLIPVLIIYCYAAIIRT 307
Db 218 IYIPALNITTCDDV-----LPEEVLGDMFNFLSLAIGVFLFPALLTASAYVLMIKT 270
Qy 308 LNA-----YDHWLWVVKASLLILVFTICFAPSNIILIIHHANYNNNDGLYFYLI 361
Db 271 LRSSAMDEHSEKKRQRAIRLIITVLAWYFICFAPSNLLLVVHYFLINTQRQSHVYALYLV 330
Qy 362 ALICGLINSCLDPFLYFLMSKT-RNHS 387
Db 331 ALCLSTLNSCIDPFVYFVSKDFRDHA 357

RESULT 2

I51667

thrombin receptor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C:Accession: U01667
R:Gerszten, R.E.; Chen, J.; Ishii, M.; Ishii, K.; Nanevics, T.; Turck, C.W.; Vu, T.H.; C
Nature 368, 648-651, 1994
A:Title: Thrombin receptor's specificity for agonist peptide is determined by its extrac
A:Reference number: U01667; MUID:94195429; PMID:8145852
A:Accession: U01667
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-420 <GER>
A:Cross-references: EMBL:U09632; NID:9495197; PIDN:AAA18498.1; PID:g495198

Query Match 27.8%; Score 599.5; DB 2; Length 420;
Best Local Similarity 34.6%; Pred. No. 1.2e-42;
Matches 128; Conservative 68; Mismatches 149; Indels 25; Gaps 8;
QY 30 LLLLLPT-----FCOSGMENDTNLAKPTLPKTR--GAPPNSPEEPFSALEGTGATITVKI 79
DB 8 LLLLLLLGAMSLCLANSSTQAKGHSNNMTIKTRIPDDSESEFEIPWDELDESGE 67
QY 80 GATITVKIKCEESASHLHVKNATMGVLTSSLSKLPAPYLLVGVVGPANAVTLMLF 139
DB 68 GSGDQAPVSRGARKPIRNRITKEAEQYLSOQLKFPVSLTVTVFVGLPLNLLAIIFL 127
QY 140 PRTSICITVY-TNLAIADPLFCVTPFKIAYHLNGNNVGFGEVLCRAITVIFGNYMC 198
DB 128 FKMVKRPVAVYMLNLAIDYFVSVLPFKIAYHLSGNDWLFPGMCRIVTAIFGNYMC 187
QY 199 SILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFLKQEYLVQ 258
DB 188 SVLLIASISVDRFLAVYPMHLSWRMSRAYMACSFIWLSIASITPLLTQTKIPR 247
QY 259 PDITTCCHDVNTCESSPFQYIFSLAFGLFPFLVLIYCYAAIIRTNAYD----- 312
DB 248 LDITTCCHVDL-LKDLKDFIYFSSFCLLFFFPFITTICYGIIIRSLSSSIENSCK 306
QY 313 -HRWLWYKASLLIIVFTICFAPSNIILIIHHANYNNYNTDGLFYLYLALCLGSLNSC 371
DB 307 KTRALF---LAVVLCVFIICGPTNVLFTH---YLQEAENEFLYFAYLSACVGSVSC 360
QY 372 LDPFLYFLMS 381
DB 361 LDPFLYFYAS 370

RESULT 3
S66518
protease-activated receptor 2 precursor - human
C:Species: Homo sapiens (man)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Nov-2000
C:Accession: S66518; S64709; G02131
R:Nystedt, S.; Emilsson, K.; Larsson, A.K.; Stroembeck, B.; Sundelin, J.
Eur. J. Biochem. 232, 84-89, 1995
A:Title: Molecular cloning and functional expression of the human prote
A:Reference number: S66518; MUID:96048032; PMID:7556175
A:Accession: S66518
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-397 <NYS>
A:Cross-references: EMBL:Z49993; NID:gl008084; PIDN:CAA90290.1; PID:g1008085
R:Boehm, S.K.; Kong, W.; Broemme, D.; Smeekens, S.P.; Anderson, D.C.; Connolly, A.; Kahn
Biochem. J. 314, 1009-1016, 1996
A:Title: Molecular cloning, expression and potential functions of the human proteinase-a
A:Reference number: S64709; MUID:96177879; PMID:8615752
A:Accession: S64709
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-137, 'A', 139-397 <BOE>
A:Cross-references: EMBL:U34038; NID:gl041728; PIDN:AAB47871.1; PID:g1041729
A:Note: the authors translated the codon GTC for residue 68 as Ile and AAC for residue 2
R:Kahn, M.L.; Coughlin, S.R.

submitted to the EMBL Data Library, September 1995
A:Reference number: H00822
A:Accession: G02131
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 29-397 <KAH>
A:Cross-references: EMBL:U36753; NID:gl208539; PIDN:AAA90957.1; PID:g1208540
C:Genetics:
A:Map position: 5q13
A:Introns: 28/1
C:Superfamily: ATP receptor P2u
F:1-36/Domain: activation peptide #status predicted <APT>
F:1-25/Domain: signal sequence #status predicted <SIG>
F:37-397/Product: proteinase-activated receptor 2 #status predicted <MAT>

Query Match 27.3%; Score 589.5; DB 2; Length 397;
Best Local Similarity 35.0%; Pred. No. 7.8e-42;
Matches 137; Conservative 65; Mismatches 138; Indels 51; Gaps 13;
QY 28 AGILLLLPTFCOSGMENDTNLAKPTLPKTRGAPPNSPEEPFSALEGTGATITVKI 87
DB 12 AAILLAASLSC-SGTIQGTRNSKGRSLIGKVDGT-----SHVTG-KGVTVETVF 59
QY 88 KCPEESASHLHVKNATMGVLTSSLSKLPAPYLLVGVVGPANAVTLMLFFRTRISCT 147
DB 60 SVDEFSAS-----VLTKLTTVFLPIVTVFVGVGLPSNGMALVFLFRKKHP 109
QY 148 TVPY-TNLAIADPLFCVTPFKIAYHLNGNNVGFGEVLCRAITVIFGNYMCILLIACI 206
DB 110 AVTYMANLALADLLSVIFPFLKIAYHGNWYIYGEALCNVLGFFGNYMCILFMTCL 169
QY 207 SINRYLAIVHPFTYRGLPKHTYALVTCGLVWATVFLYMLPFFLKQEYLVQPDITTC 266
DB 170 SVORYWVWPMGH-SRKXANIAIGSLAILWLILVITPLVYVVKQITIFALNITTC 228
QY 267 VHTTCSSSPFQY-----YFISLAFGLFPFLVLIYCYAAIIRTNLA-----YDHR 314
DB 229 V-----LPEQLLVGMFNYLSLAIGVFLPAFLTASAYVLMIRLRSSAMENSEK 281
QY 315 WLVVYKASLLIIVFTICFAPSNIILIIHHANYNNYNTDGLFYLYLALCLGSLNSC 371
DB 282 RKRAILVITVLAFLYLCITPNSLLLVH---YFLKSGQSHVYALYIVALCLSLNSC 338
QY 372 LDPFLYFLMSKT-RNHS-TAYLTKXXNDLRE 400
DB 339 IDPFVIFVSHDFRDAKALLCRSVRTVKQ 369

RESULT 4
A37912
thrombin receptor precursor - human
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 05-Nov-1999
C:Accession: A37912
R:Vu, T.K.H.; Hung, D.T.; Wheaton, V.I.; Coughlin, S.R.
Cell 64, 1057-1068, 1991
A:Title: Molecular cloning of a functional thrombin receptor reveals a novel proteolyti
A:Reference number: A37912; MUID:91188254; PMID:1672265
A:Accession: A37912
A:Molecule type: mRNA
A:Residues: 1-425 <VUA>
A:Cross-references: GB:M62424; NID:g339676; PIDN:AAA36743.1; PID:g339677
C:Genetics:
A:Gene: GDB:F2R
A:Cross-references: GDB:127737; OMIM:187930
A:Map position: 5q13-5q13
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-425/Product: thrombin receptor #status predicted <MAT>

Query Match 24.2%; Score 523; DB 2; Length 425;
Best Local Similarity 32.9%; Pred. No. 3.1e-36;
Matches 127; Conservative 76; Mismatches 153; Indels 30; Gaps 11;

```

Qy 14 GTQVWKKALIPAAAGLLLLLPTFCQSGMENDTNNAKPTLPIKTRGAPPN-SPEEPFF 72
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 2  GPRLLLLVAACFSLCGPLLSARTRARRPESKATN-ATLDPGRFLLRNPNDRKYEPFWE 57
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 73 SALEGWTGAT-ITVKKPEESASHLVKQVATMGVLTSSLSLTKLPIAYLLVFPVGV 128
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 58 DEEKNEGLTEYRLVINSKSPLOQLPAFISEDASGYLTSSWLTFLFVSVTYGVFWSL 117
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 129 PAN-AVTLMWLFRFTRISCTTVFYTNLAIAADFLFCVTLPFKIAVHLGNMVFGEVLGR 186
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 118 PLNIMAVVFIILMKVKK-PAVYMLHLATADVFLSVLPFKIISYFSGSDMQFSELGR 176
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 187 ATTVIFYGNMVCSSILLACISINRYLAIVHP--FTYRGLPKHTYALTCGLUWAVFLY 243
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 177 FVTAIFYCNMVASILLMTVISIDRELAUVYPMQSLSWRTLGRASP--TCLAIWALAING 233
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 244 MLPPFIILKQEYYLVQPDITTHCDVHNTCESSPFQLYYFISLAFFGLFIPFLIIVCYAA 303
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 234 WVPLVLRQETIQVPGNLNITTHCDVLNETLLEG-VYAYYFSAFSAVFFVPLIISTVCYVS 292
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 304 IIRTLN-----AYDHRMLVMVVKASLILVITFCFAPSNTILIIHHANY-YNNNTDGL 355
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 293 IIRCLSSAVANRKSRALF--LSAAVFCIFICFPGTNNLLIAHYSFLSHTSTTEAA 349
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 356 YFIVILALCLGSLNSCLDPFLYFLMS 381
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Qy 350 YFAYLLCVCVSSISICIDPLIYYAAS 375
Db   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT 5
A43448
Thrombin receptor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C/Accession: A43448
R/Zhong, C.; Hayzer, D.J.; Corson, M.A.; Runge, M.S.
J. Biol. Chem. 267, 16975-16979, 1992
A/Title: Molecular cloning of the rat vascular smooth muscle thrombin receptor. Evidence
A/Reference number: A43448; MUID:92381002; PMID:1324917
A/Accession: A43448
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-432 <ZHO>
A/Cross-references: GB:M81642; NID:g207465; PID:AAA42274.1; PID:g207466
A/Experimental source: RASM aortic smooth muscle cells
A/Note: sequence extracted from NCBI backbone (NCBIN:111973, NCBIPI:111974)
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match	23.4%	Score	505.5;	DB 2;	Length	432;			
Best Local Similarity	31.2%;	Pred. No.	9.1e-35;						
Matches	134;	Conservative	71;	Mismatches	163;	Indels	61;	Gaps	15;

Qy	21	KALIFAAGLULLLLPTCQSGWENDNNLAKPTLPKTI-----FRGAPNSSEPPFS	73						
	:	: :	:	:	:	:	:	:	:
Db	4	RRLULVAVGLSLCGPLSSRVPMQPESRMRYATPNNPRSPFLRNPSDETTEQFPPLG	63						
	:	:	:	:	:	:	:	:	:
Qy	74	-----ALEGTGATITVKIKCP-----EESASHLVKNATGYLTSSLSTKLI	116						
	:	:	:	:	:	:	:	:	:
Db	64	DEEKNESIPIEG--RAVLNKSFPMPPPPFIEDAS-----GYLTSPWLTLFI	112						
	:	:	:	:	:	:	:	:	:
Qy	117	PAIYLLVFVGVPANAVTLMMLFFTRTSICITTFY-TNLIADELFCVTLPFKIAYHLNG	175						
	:	:	:	:	:	:	:	:	:
Db	113	PSVYTFEIVLSPLNILAIAVVFVRMKVKKPAVVYMLHMAADVLFVSUVPFKISYFSG	172						
	:	:	:	:	:	:	:	:	:
Qy	176	NNWVFGVELCRATTVIIFYGNMYCSILLACISINRYLIAIVHP----FTYRGUPKHITYALVT	232						
	:	:	:	:	:	:	:	:	:
Db	173	TDMQFGSMCRMCFATAACYCNMYASIIMLTVISIDRFIAVVYP IQSLSWRTLGRANF---T	229						
	:	:	:	:	:	:	:	:	:
Qy	233	CGLVNAIVFLYMLPFPIKLQEYLYVQQDITTCCHVDVHTCESSSPPQLYYFISLAFFGELI	292						
	:	:	:	:	:	:	:	:	:
Db	230	CWIJWNAINMGVPELLKEQTQVPGNNITTCBVLNKETLHG-FYSYIFGAFAIGAIFFL	288						
	:	:	:	:	:	:	:	:	:

[illegible]

Query Match	23.2%;	Score	501.5;	DB	2;	Length	427;
Best Local Similarity	32.1%;	Pred. No.	2e-34;				
Matches	126;	Conservative	73;	Mismatches	140;	Indels	53;
						Gaps	13;
Qy	23	LIFAAGLLLLLPFC-----QSGMENDNNLAKTLPKTR-RGAPPNSFEFPFS	73				
Db	6	LLLVAAGLSCGPLLSRVPVRQSEMTDATVN-----PRSFPLRPNGENTFELIPLUG	59				
Qy	74	ALEGWGTGATIVKIKCPESASHLH-----VKNATMGYLTSSLSTKLIPAIYLL	122				
Db	60	DEEEKNESTL-----PEGRAIYLNKSHSPAPLAFISEDAGYLTPSWLRFIPSVYTF	113				
Qy	123	VFVVGVPAN--AVTLWMLFFETRSICTTVFYTNLAIADFLFCVTLPFKIAYHLNGNNWVF	180				
Db	114	VFVVSPLNLILAIAYFLVKMKVK--PAVYMLHLAMADVLFVSVLPLKISYFSGSDMOF	172				
Qy	181	GEVLCRATTVIFYGNMYSILLACISINRYLAIVHP--FTYRGLPKHTYALVTCGLVM	237				
Db	173	GSGMCRFATAFYCNMYASIMLMTVISIDRELAVVYPIQSISWRTLGRANF---TCLVIV	229				
Qy	238	ATVELYMLPPFILKOEYVLVOPDITTHCDVHNTCESSPPOLYFISLAFFGFLIPFVLI	297				
Db	230	VMAIMGVVPULLKEQTRVPGLNITTHCDVLNETLLOG-FYSYIYFSAFSAVFPLVPLIIS	288				
Qy	298	IYCYAAIIRTNLAYD-----HRLMWVKASLLILVITFCFAPSNIILITHHANYYN	350				
Db	289	TICWMSIIRCUSSSVANRKSRALF---JSAAVCFVIVCFGFTNLLIMHYLLLLSDS	345				
Qy	351	-NTDGLYFIYLIALCLGSLNSCLDPFLYFLMS	381				
Db	346	PATEKAYFAYLLCCVCSVSSVCIDPLIYVYAS	377				

RESULT 7
S33733
G protein-coupled receptor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: S33733
R:Webb, T.E.; Simon, J.; Kriehke, B.J.; Bateson, A.N.; Smart, T.G.; King, B.
FEBS Lett. 324, 219-225, 1993
A:Title: Cloning and functional expression of a brain G-protein-coupled APT
A:Reference number: S33733; MUID:93285340; PMID:8508924

A;Accession: G33733
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-362 <WEB>
A;Cross-references: EMBL:X73268; NID:g395084; PIDN:CAA51716.1; PID:g395085
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 19.7%; Score 424; DB 2; Length 362;
Best Local Similarity 28.9%; Pred. No. 5.1e-26;
Matches 101; Conservative 59; Mismatches 143; Indels 46; Gaps 9;

Qy 77 GWTGATITVKIKPEESASHLRVKNATMGYLTSSLSKTLPIAYILVVFVGVGPANAVTLW 136
Db 21 GWAAGNATTKSLTKTGQFY-----LPTVILVFIITGFLGNSVAIW 63

Qy 137 MLFRTSRIC-TTVFYNNLAIADELFCVTLDPFKIAYHLNGNNWVGEVLCRATTVIYGN 195
Db 64 MFVFMRPWSGISVYMFENALADEFLYTLPALIFYFNKTDWTFGDMCKLQRFIFHN 123

Qy 196 MYCSILLACISINRYLAIVHPFYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEY 255
Db 124 LYSGILFICSVHRYTGWVHPLKSLGLRKKNAVYVSSVWALVAVIAPIL-----FY 178

Qy 256 ----LVOPDITCTHDVHNTCESSPFFQLYYFISLAFFGLIPFVLIIVCYAAIIRTLNAY 311
Db 179 SGTGVRRNKITCYD--TTADEYLSYFVYSMCTTVFMFCIPFVILGCGYLIKVALIYK 236

Qy 312 D-----HRWLVKASLLILVITTCFAPSNNILIH-HANYYYN-----TDGLYF 357
Db 237 DLDSPLRKSIYLV--IIVLVFAVSYPFLVHVKTNLRLARDFTQTPQCAFNKVYA 293

Qy 358 IYLALCIGSINSLDPELYELMSKTRNHSYATLTXXNDLREQQPSQ 406
Db 294 TYQVTRGLASLNSCDPILYFLAGDTFRRRLSRATRSRSEPNVQSK 342

RESULT 8

B45680
G protein-coupled peptide receptor EBI 2 - human
C;Species: Homo sapiens (man)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: B45680
R;Birkenbach, M.; Josefsen, K.; Yamamachili, R.; Lenoir, G.; Kieff, E.
J. Virol. 67, 2209-2220, 1993
A;Title: Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled R
A;Reference number: A45680; MUID:93188173; PMID:8383238
A;Accession: B45680
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-361 <BI>
A;Cross-references: GB:L08177; NID:g292056; PIDN:AAA35924.1; PID:g292057
A;Experimental source: B-lymphocytes
A;Note: sequence extracted from NCBI backbone (NCBIN:127096, NCBIPI:127097)
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 19.0%; Score 410.5; DB 2; Length 361;
Best Local Similarity 30.2%; Pred. No. 6.8e-27;
Matches 99; Conservative 58; Mismatches 114; Indels 57; Gaps 10;

Qy 106 YLTSSLSKLPATYLLVVFVGVGPANAVTLWMLFFRTRSI-CTTVFYNNLAIADELFCVCT 164
Db 24 YAHSTARIWVPLHSLYFLIGLVGNLLALVIVQNRKKNSTLYSTNLVISILTTA 83

Qy 165 LPFKIAYHLNGNNWVGEVLCRATTVIYGNMYCSILLACISINRYLAIVHPFYRGLP 224
Db 84 LPTRIAVYAMGFDWRIGDALCRITALVFINTYAGVNFMTCLSIDRFTIAVHPLRYNKK 143

Qy 225 KHTVALVTCGLVWATVFLYMLPFFI---LKQEYLVQPDITCTHDVHNTCESS-PPQLY 280
Db 144 RIEHAGVCIFWILVFAQTLLINPMNSKOE-----AERTICYEPNFETKSLPIL- 197

Qy 281 YFISLAFFGFLIPFVLIIVCYAAII-----RTLNAYDHRWLWVYK 320
Db 198 --LGACFIGVPLIILICYSQICKLPRTAQNPLTEKSGVKKALNTI----- 246

Qy 321 ASLLIVIFITCAPSNILIIH-----HANYYYNNTDGLYFIYL-IALCLGSLNSCLD 373
Db 247 --ILIIIVFVLCFTPYHVAIOHMIKKLRFESNLECSQRHSFQISLFTVCLMNFNCMD 304

Qy 374 PKLYELMSKTRNHSYATLTXXNDLREQ 401
Db 305 PFIYFFACK-----GYKKRVMRLKQ 326

RESULT 9
I50241
G protein-coupled receptor 6H1 - chicken
N;Alternate names: purinoceptor 6H1
C;Species: Gallus gallus (chicken)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 02-Jun-2000
C;Accession: I50241; JC4618
R;Kaplan, M.H.; Smith, D.I.; Sundick, R.S.
J. Immunol. 151, 628-636, 1993
A;Title: Identification of a G protein coupled receptor induced in activated T cells.
A;Reference number: I50241; MUID:93329058; PMID:8393036
A;Accession: I50241
A;Status: preliminary; translated from GB/EMBL/DBD
A;Molecule type: mRNA
A;Residues: 1-308 <KAP>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
R;Webb, T.E.; Kaplan, M.G.; Barnard, E.A.
Biochem. Biophys. Res. Commun. 219, 105-110, 1996
A;Title: Identification of 6H1 as a p2Y purinoceptor: p2Y5.
A;Reference number: JC4618; MUID:96190677; PMID:8619790
A;Accession: JC4618
A;Molecule type: mRNA
A;Residues: 1-308 <WEB>
A;Cross-references: GB:L06109; NID:g304383; PIDN:AAB06587.1; PID:g304384
A;Experimental source: T-cells
A;Comment: This receptor plays a role in T-cell activation.
C;Genetics:
C;Gene: p2Y5
C;Superfamily: ATP receptor P2u
C;Keywords: G protein-coupled receptor; transmembrane protein
F;15-40/Domain: transmembrane #status predicted <TM1>
F;51-74/Domain: transmembrane #status predicted <TM2>
F;83-109/Domain: transmembrane #status predicted <TM3>
F;133-153/Domain: transmembrane #status predicted <TM4>
F;177-201/Domain: transmembrane #status predicted <TM5>
F;227-248/Domain: transmembrane #status predicted <TM6>
F;269-292/Domain: transmembrane #status predicted <TM7>

Query Match 18.8%; Score 406.5; DB 2; Length 308;
Best Local Similarity 33.5%; Pred. No. 1.3e-26;
Matches 108; Conservative 52; Mismatches 127; Indels 35; Gaps 11;

Qy 107 LTSLSLTK-----LIPAIYLLVVFVGVGPANAVTLWMLFF--RTRSICTTVFYNNLAIA 158
Db 2 VSSNCSTEDSEFKTLTYCCVFSVMFVLGLIANCAIVIFFTLKVNRN-TTYMLNLAISD 60

Qy 159 FLFCVTLPPFKIAYHLNGNNWVGEVLCRATTVIYGNMYCSILLACISINRYLAIVHPF 218
Db 61 LLFVFTLPFRI-YFVVRNWPFGDVLCKISVTLFTYNNMGYSILFLTCISVDRLAIVHPF 119

Qy 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYLVQPDITCTHDVHNTCESS-SPP 277
Db 120 RSKTLRTRNARIVCAVAVITVLGASTPASFPSTQNNNTQRTQCFE--NFPESTWKTY 177

Qy 278 QLQYFISLAFFGFLIPFVLIIVCYAAIIRTLN---AYDHRWLWVYKASLLI---LVIFTI 331
Db 178 LSRIVFIEIVGFPIPLINVTCTWVLRNLKPLTLNRNKLKSKKVLKMFVHLVIFCF 237

Qy 332 CFAPSNILIIH---HANYYYNNT--DGLYFIYLIACLGSLNSCLDPPFLYFLMSKTRNH 386

Db 238 CFVFNITLILSLMRTQWNCVSVTAIRTMYPVTLCLIAVSNCCFDDIVVYFTSDTN-- 295

Qy 387 STAYLTXXXNDLRQGGQPSQRT 408

Db 296 -----SELDDKKQVHQNT 308

RESULT 10

JC4737

G protein-coupled receptor P2Y1 - human

N;Alternate names: P2Y1 purinergic receptor; P2Y1 purinoreceptor

C;Species: Homo sapiens (man)

C;Date: 10-May-1996 #sequence revision 16-Aug-1996 #text_change 17-Nov-2000

C;Accession: JC4737; JC4615; S54253

R;Janssens, R.; Communi, D.; Pirotton, S.; Samson, M.; Parmentier, M.; Boeynaems, J.M. Biochem. Biophys. Res. Commun. 221, 588-593, 1996

A;Title: Cloning and tissue distribution of the human P2Y1 receptor.

A;Reference number: JC4737; MUID:96205320; PMID:8630005

A;Accession: JC4737

A;Molecule type: DNA

A;Residues: 1-373 <JAN>

A;Cross-references: GB:S81950; NID:g1839438; PIDN:AAA47091.1; PID:g1839439

R;Ayyanathan, K.; Webb, T.E.; Sandhu, A.K.; Athwal, R.S.; Barnard, E.A.; Kunapuli, S.P. Biochem. Biophys. Res. Commun. 218, 783-788, 1996

A;Title: Cloning and chromosomal localization of the human P2Y1 purinoreceptor.

A;Reference number: JC4615; MUID:96158962; PMID:8579591

A;Accession: JC4615

A;Molecule type: mRNA

A;Residues: 1-373 <AY>

A;Cross-references: GB:U42029; NID:g1147730; PIDN:AAA97872.1; PID:g1147731

A;Experimental source: erythro leukemia cells

R;Leon, C.; Vial, C.; Cazenave, J.; Gachet, C. submitted to the EMBL Data Library, May 1995

A;Description: Cloning of a human putative P2Y receptor.

A;Reference number: S54253

A;Accession: S54253

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-137,139-373 <LEO>

A;Cross-references: EMBL:Z49205; NID:g798835; PIDN:CAA89066.1; PID:g798836

C;Comment: This receptor belongs to a family of G protein-coupled receptors. It responds to ATP.

C;Genetics:

A;Gene: P2Y1; GDB:P2Y1

A;Cross-references: GDB:677125; OMIM:601167

A;Map position: 3pter-3qter

C;Superfamily: ATP receptor P2u

C;Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

F;88-111/Domain: transmembrane #status predicted <TM1>

F;124-152/Domain: transmembrane #status predicted <TM2>

F;171-191/Domain: transmembrane #status predicted <TM3>

F;214-237/Domain: transmembrane #status predicted <TM4>

F;261-282/Domain: transmembrane #status predicted <TM5>

F;305-328/Domain: transmembrane #status predicted <TM6>

F;11,27,113,197/Binding site: carboxydrate (Asn) (covalent) #status predicted

F;258,336/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

F;330,339/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

F;343/Binding site: phosphate (Ser) (covalent) (by protein kinase C and calmodulin-depen

Query Match 18.5%; Score 399; DB 2; Length 373;

Best Local Similarity 29.7%; Pred. No. 6.5e-26;

Matches 113; Conservative 66; Mismatches 143; Indels 58; Gaps 14;

Qy 52 PTLPIKT---FRGAPPNSFEFFPSALEGWTGATITVKIKCPESASHLVKNATMGYLT 108

Db 7 PAVNGTDAFLACPGSS-----WGNSTV-----ASTAAVSSFKCALT 45

Qy 109 -SSLSKTLIPALVLLVFWGVGPANAVTLWMLFFETRSIC-TTVFTYTLAIADFLFCVTLP 166

Db 46 KTGQFQFYLLPAVYLIVFIIGLGNVAIWMFVFMKPGWSGISVYMFNLALADFLVLTLP 105

Qy 167 FKIAHLNGNNVWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFTYRGLPKH 226

Db 106 ALIPYFNKTDWIFGDAMCKLQRFIFHNLYGSILFTCTISAHRYSGVVPLKSLGLRKK 165

Qy 227 TYALVTCGLVWATVFLYMLPEFFILKQEYLYVQPDIT-TCHDVHNTCESSPPFQLYYFI-- 283

Db 166 KNAICISLVWLIWVAISP--ILFYSGTGVKNKTTITCYDT-----TSDEYLSYFIYS 218

Qy 284 ---SLAPFGFLIPVLIICYAAIIRTLNAYD-----HRWLWYVKASLLILVIFTICF 333

Db 219 MCTTVAM--FCVPLVLILGCVGLIVRALIYKLDLNSPLRRKSIYLV---IIVLTVFAVS 273

Qy 334 APSNIIILIH-HANYNN-----TDGLYFIYLIJALCLGSLNSCLDPFLYFLMSKTRNH 386

Db 274 IPFHWMTKMLRLRLDQTPAMCAFNDRVATYQVTRGLASLNSCVDPILYFLAGDTFRR 333

Qy 387 STAYLTXXXNDLRQGGQPSQ 406

Db 334 RLSRAIRKASRRSEANLQSK 353

RESULT 11

JC4162

P2Y receptor - bovine

C;Species: Bos primigenius taurus (cattle)

C;Date: 12-Oct-1995 #sequence revision 10-Nov-1995 #text_change 24-Sep-1999

C;Accession: JC4162

R;Henderson, D.J.; Elliot, D.G.; Smith, G.M.; Webb, T.E.; Dainty, I.A. Biochem. Biophys. Res. Commun. 212, 648-656, 1995

A;Title: Cloning and characterisation of a bovine P2Y receptor.

A;Reference number: JC4162; MUID:95352058; PMID:7626079

A;Accession: JC4162

A;Molecule type: mRNA

A;Residues: 1-373 <HEN>

A;Cross-references: EMBL:X87628; NID:g1032484; PIDN:CAA60958.1; PID:g1032485

A;Experimental source: aortic endothelial cell

C;Genetics:

A;Gene: bovp2y

C;Superfamily: ATP receptor P2u

C;Keywords: glycoprotein; phosphoprotein; receptor; transmembrane protein

F;52-77/Domain: transmembrane #status predicted <TM1>

F;88-111/Domain: transmembrane #status predicted <TM2>

F;124-150/Domain: transmembrane #status predicted <TM3>

F;171-191/Domain: transmembrane #status predicted <TM4>

F;214-237/Domain: transmembrane #status predicted <TM5>

F;261-282/Domain: transmembrane #status predicted <TM6>

F;305-328/Domain: transmembrane #status predicted <TM7>

F;11,27,113,197/Binding site: carboxydrate (Asn) (covalent) #status predicted

F;258/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 18.5%; Score 398; DB 2; Length 373;

Best Local Similarity 28.5%; Pred. No. 7.8e-26;

Matches 111; Conservative 66; Mismatches 136; Indels 76; Gaps 15;

Qy 45 DTNNLAKPTLPIKTRGAPPNSFEFFPSALEGWTGATI---TVKIKCPESASHLVKN 102

Db 14 DTAFLADP-----GSP-----WGNSTVTSTAASPFKCA----- 43

Qy 103 TMGYLT-SSLSKTLIPALVLLVFWGVGPANAVTLWMLFFETRSIC-TTVFTYTLAIADFL 160

Db 44 ---LTKTGQFYLLPAVYLIVFIIGLGNVAIWMFVFMKPGWSGISVYMFNLALADFL 99

Qy 161 FCVTLPEFKIAYHLNGNNVWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFTY 220

Db 100 YVLTLPALIFYFNKTDWIFGDAMCKLQRFIFHNLYGSILFTCTISAHRYSGVVPLKS 159

Qy 221 RGLPKHTYALVTCGLVWATVFLYMLPEFFILKQBY---LVQPDITTCCHDVHNTCESSSP 276

Db 160 LGRKKKNNAVYISVLVWLIWVAISP-----FYSGTGIRKNTKITCYDT-----TSDE 209

Qy 277 FQLYYFI-----SLAPFGFLIPVLIICYAAIIRTLNAYD-----HRWLWYVKASLL 324

Db 210 YLSYFIYSMCTTVAM--FCVPLVLILGCVGLIVRALIYKLDLNSPLRRKSIYLV---II 264

Qy 325 ILVIFTICFAPSNIILIH-HANYNN-----TDGLYFIYLIJALCLGSLNSCLDPFLY 377

```
Db 265 VLTVFAYSYPPFHVMKTNLRARLDFTQPEMCAENDRVATYQVTRGLASINSCVDPIY 324
Qy 378 FLMSKTRNHSYALTXXNDLREGQPSQ 406
Db 325 FLAGDTFRRLSRATRKASRRSEANLQSK 353

RESULT 12
JC5549
heptahelical P2Y5-like receptor - human
C:Species: Homo sapiens (man)
C:Date: 02-Sep-1997 #sequence_revision 05-Sep-1997 #text_change 24-Sep-1999
C:Accession: JC5549
R:Janssens, R.; Boeynaems, J.M.; Godart, M.; Communi, D.
Biochem. Biophys. Res. Commun. 236, 106-112, 1997
A:Title: Cloning of a human heptahelical receptor closely related to the P2Y5 receptor.
A:Reference number: JC5549; MUID:97366605; PMID:9223435
A:Accession: JC5549
A:Molecule type: DNA
A:Residues: 1-370 <JAN>
A:Cross-references: DDBJ:AF005419; NID:g2240034; PIDN:AAB66322.1; PID:g2240035
C:Superfamily: ATP receptor P2u

Query Match 17.8%; Score 384.5; DB 2; Length 370;
Best Local Similarity 30.4%; Pred. No. 1.1e-24;
Matches 95; Conservative 60; Mismatches 121; Indels 37; Gaps 10;

Qy 91 BESASHLHVK--NATMG---YLTSSLSTKLIPAIYLLVFGVGPANAVTLMMLFFRTRSI 145
Db 12 QDSNSLRPRLGATNATNCIVDSFKYLNGLAVSYVFLIGLITNSVLSFLVFCFRMKMR 71

Qy 146 C-TTVFVYNLAIAFLFCVTLPPFKIAYHLNNGNWFGEVLCRAITVIFGYNMVCSILLLA 204
Db 72 SETAIFITNLAVSDLLFVCTLPFKIFYNFN-RHWPFGDTLCKISGTAFATNIYGSMLFLT 130

Qy 205 CISINRYLAIVHPETRYGLPKHTYALVTCGLAVATVFLYMLPFFILKQEYVLQDPDITTC 264
Db 131 CISVDRFVAIVYPPRSRTIRTRNSAIVCAGWILVLSGGISALF-----STTNV 181

Qy 265 HDVHNTCESSSPFFQY--YFISLAFF----GFLIPFVLIICYAAIIR-----TLNAYD 312
Db 182 NNATTTCPEGLSKRWKTYLSKITIFIEVVGFIPIILNVSCSSVWLRLRKPKATLSQIG 241

Qy 313 HRWLWYKVASLLILVIFTICFAPSNILIIHH-----ANYNNYNDGLFYFLIALCL 365
Db 242 TNKKKVLKMTVHMAVFVCEVPYNSVLFYALVRSQAITNCFLERFAKI--MYPITLCL 299

Qy 366 GSLNSCLDPFLYF 378
Db 300 ATLNCDFDFIYY 312

RESULT 13
S68679
G protein-coupled receptor - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68679
R:Stam, N.J.; Klomp, J.; van de Heuvel, M.; Olijve, W.
FEBS Lett. 384, 260-264, 1996
A:Title: Molecular cloning and characterization of a novel orphan receptor (P(2P)) expre
A:Reference number: S68679; MUID:96197801; PMID:8617367
A:Accession: S68679
A:Molecule type: DNA
A:Residues: 1-365 <STA>
A:Cross-references: EMBL:X96597; NID:g1296631; PIDN:CAA65415.1; PID:g1296632
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor

Query Match 17.7%; Score 382; DB 2; Length 365;
Best Local Similarity 31.8%; Pred. No. 1.7e-24;
```

```
Matches 100; Conservative 47; Mismatches 127; Indels 40; Gaps 10;

Qy 115 LIPAIYLLVFGVGPANAVTLMMLFFRTRSICTTVEYT-NLAIADEFLFCVTLRPFKAYHL 173
Db 37 LLEVSVAVVFVGLGNAPLMLFIFRLRPWDATYMFHLAUSDLYLSLSTPLIYYA 96

Qy 174 NGNNVFGVLCRAITVIFGYNMVCSILLIACISINRYLAIVHPETRYGLPKHTYALVTC 233
Db 97 ANHNWPFGEI CKFVRFVFNLYCSVLFCTCISVHYLGICHPLRALRWGRPRLAGLLC 156

Qy 234 GLVMATVFLYMLP--FFILKQEYVLQDPDITTCVDHNTCESSSPFQLYFISLAFPGFL 291
Db 157 LAWLAVAGCLVNLPEVTTSN---KGTIVLCHDTRPEE---FDHYVHFSSAVMGLL 208

Qy 292 --IPFVLIICYAAIIRTL-----NAYDHRWLWYKVASLLIIVITICFAPSNIILIH 343
Db 209 FGVPCLVTLVCYGLMARRYLQPLPGSAQSSRSRLRTIAVLTVFVCFVPHITRTIY 268

Qy 344 H-----ANYNNYNDGLFYFLIALCLGSLNSCLDPFLYFLMSKTRNHSYALTXXND 397
Db 269 YLARLLEADCRVLNI--VNVVYKVTPLASANSCLDPVLYLLTG-----DKYRRQ 316

Qy 398 LRE---QGQPSQRT 408
Db 317 LRQLCGGKGPQRT 330

RESULT 14
T09508
intron 17 purinergic receptor P2Y5 - human
N:Alternate names: G-protein coupled receptor
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09508
R:Bohm, S.K.; Trumpp, A.; Khitlin, L.M.; Kong, W.; Payan, D.G.; Bunnett, N.W.
submitted to the EMBL Data Library, April 1997
A:Description: The human purinergic receptor P2Y5 is encoded in intron 17 of the retino
A:Reference number: Z16705
A:Accession: T09508
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-344 <BOH>
A:Cross-references: EMBL:AF000546; NID:g2232068; PID:g2232069
C:Genetics:
A:Map position: 13
C:Superfamily: ATP receptor P2u
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 17.4%; Score 376; DB 2; Length 344;
Best Local Similarity 30.9%; Pred. No. 5e-24;
Matches 96; Conservative 55; Mismatches 122; Indels 38; Gaps 10;

Qy 101 NATMGYLTSSLSTKLIPAIYLLVFGVGPANAVTLMMLFFRTRSICTTVEYTN 153
Db 5 NSSHCFVNDYFKVTLYGCMFVFLGLVSNCAIYIF-----ICVLKVRNETTYMIN 58

Qy 154 LAIADEFLFCVTLRPFKAYHLNNGNWFGEVLCRAITVIFGYNMVCSILLIACISINRYLA 213
Db 59 LAMSDDLFLVFTLPFRI--FYFTRNWPFGDLLCKISVMLFYTNMYGSLFTICISVDRFLA 117

Qy 214 IVHPETRYGLPKHTYALVTCGLVMATVFLYMLPFFILKQEYVLQDPDITTCVDHNTCES 273
Db 118 IVFPFKSKTLRTKRNKIVCTGWLTVIGSAP-----AVFQSTHSGQNNASEACFE 170

Qy 274 SSPFQLY--YFISLAFF----GFLIPFVLIICYAAIIRTLN---AYDHRWLWYKVASLL 324
Db 171 NFPEATWKTYSRIVIFIEIVGFFIPLNLNVTCSMWLKTLPKPVTLRSRKINKTKVKM 230

Qy 325 I---LVFTTCFAPSNIILIH---HANYNNYNT--DGLFYFLIALCLGSLNSCLDPFL 376
Db 231 IFVHLIIFCFVFPYINILYLSLVRTQTQTFVNCVAAVRTMYPITLICIAVSNCCFDPV 290

Qy 377 YFLMSKTRNHS 387
```


Db 291 YFTSDTQNS 301

RESULT 15

QJ1231

Interleukin-8 receptor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Nov-1999

C:Accession: JQ1231; A46483

R:Backmann, M.P.; Munger, W.E.; Kozlosky, C.; VandenBos, T.; Price, V.; Lyman, S.; Gerard

Biochem. Biophys. Res. Commun. 179, 784-789, 1991

A:Title: Molecular characterization of the interleukin-8 receptor.

A:Reference number: JQ1231; MUID:91378994; PMID:1898400

A:Accession: JQ1231

A:Molecule type: DNA

A:Residues: 1-355 <BEC>

A:Cross-references: GB:M74240; NID:g165438; PIDN:AAA31375.1; PID:g165439

R:Lee, J.; Kuang, W.J.; Rice, G.C.; Wood, W.I.

J. Immunol. 148, 1261-1264, 1992

A:Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.

A:Reference number: A46483; MUID:92148149; PMID:1737938

A:Accession: A46483

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-355 <LEE>

A:Cross-references: GB:M82873; NID:g165440; PIDN:AAA31376.1; PID:g165441

A:Experimental source: neutrophils

A>Note: sequence extracted from NCHI backbone (NCBIN:81526, NCBIP:81530)

C:Superfamily: vertebrate rhodopsin

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.9%; Score 363.5; DB 2; Length 355;

Best Local Similarity 29.8%; Pred. No. 5.8e-23;

Matches 100; Conservative 66; Mismatches 101; Indels 69; Gaps 14;

Qy 107 LTSSLSTKLIPAIYLLVFWGVGVPANA-VTLMLPFRTRISCTTVFYTNLAIADELFVCVTL 165

Db 38 VTQTLNKYVVVVVIALVFLLSLGLNSLVMLVILYSGNSRSTVDVYLNLAADLLFALTM 97

Qy 166 PFKATYHLNGNNWVGEVLCRATTVFYGNMVCISILLIACISINRYLAIVHPFTYRGPLK 225

Db 98 PIWAVSKEG--WIFGTPCKVSLVKEVNFYSIGILLACISVDRYLAIVHATRLTKR 155

Qy 226 HTYALVTCGLWAVVFLVMLPFFFLKQEYLVQDPITT--CHD--VNHTCESSPPPOLY 281

Db 156 HLKVFICLG-INALSLILSLPFLFRQVF---SPNNSPVCYEDLGHNTAK----WRVVL 207

Qy 282 FISLAFTGFLIPFVLIYCYAAIIRTL----NAYDHRWLWYVKASLLILVIFTCFAPS 337

Db 208 RILPHTFGFILPLVLMVLCYGFRTLTLFQAHMGQKRAMRVIFA---VWLIFLCLWLPYN 264

Qy 338 ILLIHHANNYYNNTDGLYFYLI-----AL-----CLGSLNSCLDPFLYF--- 378

Db 265 LVLL-----ADTLMRTHVIOETCQRNDIDRALDATEILGLFLHSLNPIIYAFIG 314

Qy 379 -----LMSK-----TRNHSTAYLTXXN 396

Db 315 QNFRNGFLKMLAARGLSISKEFLTRHRTSYTSSSTN 350

RESULT 16

A47556

ATP receptor P2u - mouse

C:Species: Mus musculus (house mouse)

C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999

C:Accession: A47556

R:Lustig, K.D.; Shiao, A.K.; Brake, A.J.; Julius, D.

Proc. Natl. Acad. Sci. U.S.A. 90, 5113-5117, 1993

A:Title: Expression cloning of an ATP receptor from mouse neuroblastoma cells.

A:Reference number: A47556; MUID:93281707; PMID:7685114

A:Accession: A47556

A>Status: Preliminary

Query Match 16.5%; Score 356.5; DB 2; Length 362;

Best Local Similarity 30.3%; Pred. No. 2.3e-22;

Matches 90; Conservative 53; Mismatches 132; Indels 34; Gaps 9;

Qy 115 LIPAIYLLVFWGVGVPANAVTLMLPFRTRS-ICTTVFYTNLAIADELFVCVTLPPKIAVHL 173

Db 35 LLPVSYGVCVGLCLNNVALYIFLCRLKTNWASTTTFMFLAVSDSLYASLASLLVLYYA 94

Qy 174 NGNNWVGEVLCRATTVFYGNMVCISILLIACISINRYLAIVHPFTYRGPLKHYALVTC 233

Db 95 RGDHWPSTVLCKLVRLFFYTNLYCSILFLTCISVHRCGLVLRPLHSLRWGRARYARVA 154

Qy 234 GLVWATVFLVMLP--PFILKQEYLVQDPITTCCHDVHNTCESSPPPOLYFISLAFFGFL 291

Db 155 AVWVVLVLAQAPVLPVFTTS----VRGTRITCHDT-SARELFSHFVAYSSVMLGLL-FA 208

Qy 292 IPFVLIYCYAAIIR-----TLNAYDHRWLWYVKASLLILVIFTCFAPSNIILIIH 343

Db 209 VPFSVLVLCVLMARLLKPAYGTGGLPRAKRSVRTIALVLAVFALCFPLPHVHTLY 268

Qy 344 HANNYYNNTD-----GLYFIYLIACLGSLNSCLDPFLYFLMSKTRNHSTAYLTXXN 396

Db 269 ---YSPRSLDLSCHTLNAINWAKYITRPLASANSCLDPVLYFLAGOR-----LVRFAR 318

Qy 397 DLREQQGPS 405

Db 319 DAKPPTPT 327

RESULT 17

JN0694

angiotensin II receptor - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000

C:Accession: JN0694; S47627

R:Ji, H.; Sandberg, K.; Zhang, Y.; Catt, K.J.

Biochem. Biophys. Res. Commun. 194, 756-762, 1993

A:Title: Molecular cloning, sequencing and functional expression of an amphibian angiotensin II receptor.

A:Reference number: JN0694; MUID:93343933; PMID:7688227

A:Accession: JN0694

A:Molecule type: mRNA

A:Residues: 1-362 <JIH>

A:Cross-references: GB:L16463; NID:g387890; PIDN:AAA49647.1; PID:g387891

R:Nishimatsu, S.; Koyasu, N.; Sugaya, T.; Ohnishi, J.; Yamagishi, T.; Murakami, K.; Miy

Biochim. Biophys. Acta 1218, 401-407, 1994

A:Title: Isolation and characterization of two alternatively spliced complementary DNAs

A:Reference number: S47627; MUID:94325348; PMID:7519446

A:Accession: S47627

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-362 <NIS>

A:Cross-references: GB:S73274; NID:g625009; PIDN:AAC60749.1; PID:g625010

C:Superfamily: vertebrate rhodopsin

C:Keywords: glycoprotein; phosphoprotein; transmembrane protein

F:34-54/Domain: transmembrane #status predicted <TM1>

F:66-86/Domain: transmembrane #status predicted <TM2>

F:105-125/Domain: transmembrane #status predicted <TM3>

F:146-166/Domain: transmembrane #status predicted <TM4>

F:192-212/Domain: transmembrane #status predicted <TM5>

F:241-260/Domain: transmembrane #status predicted <TM6>

F:285-305/Domain: transmembrane #status predicted <TM7>

F:318-177/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:222,343/Binding site: phosphate (Thr) (covalent) #status predicted

F:347/Binding site: phosphate (Ser) (covalent) #status predicted

A:Molecule type: mRNA

A:Residues: 1-373 <LUS>

A:Cross-references: GB:L14751; NID:g309457; PIDN:AAA39871.1; PID:g309458

C:Superfamily: ATP receptor P2u

C:Keywords: transmembrane protein

Query Match 16.7%; Score 361; DB 2; Length 373;

Best Local Similarity 29.1%; Pred. No. 9.8e-23;

Matches 90; Conservative 53; Mismatches 132; Indels 34; Gaps 9;

Qy 115 LIPAIYLLVFWGVGVPANAVTLMLPFRTRS-ICTTVFYTNLAIADELFVCVTLPPKIAVHL 173

Db 35 LLPVSYGVCVGLCLNNVALYIFLCRLKTNWASTTTFMFLAVSDSLYASLASLLVLYYA 94

Qy 174 NGNNWVGEVLCRATTVFYGNMVCISILLIACISINRYLAIVHPFTYRGPLKHYALVTC 233

Db 95 RGDHWPSTVLCKLVRLFFYTNLYCSILFLTCISVHRCGLVLRPLHSLRWGRARYARVA 154

Qy 234 GLVWATVFLVMLP--PFILKQEYLVQDPITTCCHDVHNTCESSPPPOLYFISLAFFGFL 291

Db 155 AVWVVLVLAQAPVLPVFTTS----VRGTRITCHDT-SARELFSHFVAYSSVMLGLL-FA 208

Qy 292 IPFVLIYCYAAIIR-----TLNAYDHRWLWYVKASLLILVIFTCFAPSNIILIIH 343

Db 209 VPFSVLVLCVLMARLLKPAYGTGGLPRAKRSVRTIALVLAVFALCFPLPHVHTLY 268

Qy 344 HANNYYNNTD-----GLYFIYLIACLGSLNSCLDPFLYFLMSKTRNHSTAYLTXXN 396

Db 269 ---YSPRSLDLSCHTLNAINWAKYITRPLASANSCLDPVLYFLAGOR-----LVRFAR 318

Qy 397 DLREQQGPS 405

Db 319 DAKPPTPT 327

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-Aug-1999

R/Accession: S42096
C/Gobl, A.E.; Wang, S.; Zhou, Y.; Oeberg, K.
submitted to the EMBL Data Library, February 1994
A/Description: Molecular cloning of the rat IL8 receptor.
A/Reference number: S42096
A/Accession: S42096
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-356 <GOB>
A/Cross-references: EMBL:X77797
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 16.4%; Score 353; DB 2; Length 356;
Best Local Similarity 30.3%; Pred.No. 4.4e-22;
Matches 95; Conservative 59; Mismatches 116; Indels 44; Gaps 11;

QY 116 IPAIYLIVFVVGVPANAVTLWMLFFRTRSTC--TTVFYTNLAIADFLPCVTLPFKIAYHL 173
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
50 VVIYVLVLTSLVGNLSVMVLVILY-NRSTCSVDVLLNLAIDLAFALTLPVWAASKV 108
QY 174 NGNNWFGEVLCRAITTVIFGNYMYSILLACISINRYLAIVHPFTVRGLPKHTYALVTC 233
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
109 NG-WIFSGLCKVFSFLQBITFVSVVLLACTSMRDYLAIVHATSTLIQKRHLVKV-C 165
QY 234 GLVNAVFLYMLPFIKLBQBYLVQPDIITCHDVHNTHCESSSPQLYYFISLAFPGFLIP 293
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
166 ITWFLSLVLSPLFILRTT-VKANPSWTWCYE--NIGNNTSKRWVRLRPQTGYGLFP 222
QY 294 FVLIIYCVAAILRTL----NAYDHWLWYVKASILLIVFTICFAPSNIILLIHANNYYY 349
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
223 LLIMLFCYGTGRTLFLKAHQKHARMVIFA--VVLFELLCWLPPYNILVF----- 271
QY 350 NNTDGLFYIYLI-----ALCLGSLSCLDPFLFYFLMSKTRNHS-----TAYL 391
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
272 --TDLMRTKLKETCERQNEINKASEILGHSLNPPIIVAFIGOKFRHGLLKIMANYG 329
QY 392 TKXNDLRBOGQPS 405
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
330 LVSKEFLAKEGRPS 343

RESULT 21
I55450
G protein-coupled P2 receptor - rat
C/Species: Rattus norvegicus (Norway rat)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 19-May-2000
C/Accession: I55450
R/Chang, K.; Hanaoka, K.; Kumada, M.; Takuwa, Y.
J. Biol. Chem. 270, 26152-26158, 1995
A/Title: Molecular cloning and functional analysis of a novel P2 nucleotide receptor.
A/Reference number: I55450; MUID:96064682; PMID:7592819
A/Accession: I55450
A/Status: preliminary; translated from GB/EMBL/DBDJ
A/Molecule type: mRNA
A/Residues: 1-328 <RES>
A/Cross-references: GB:D63665; NID:g1066007; PIDN:BAA09816.1; PID:g1066008
C/Superfamily: ATP receptor P2u
C/Keywords: G protein-coupled receptor

Query Match 16.3%; Score 352; DB 2; Length 328;
Best Local Similarity 31.4%; Pred.No. 4.9e-22;
Matches 96; Conservative 42; Mismatches 126; Indels 42; Gaps 10;

QY 115 LPFAIYLLVFWGVGPANAVTLWMLFFRTRSTCTTFVT--NLAIADFLCVTLTPFKIAYHL 173
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
28 LLPVSVSWLVWGVLNVVCVIAQICASSRRTLTRSADVTLNALADLLYACSLPLIINYA 87
QY 174 NGNNWFGEVLCRAITTVIFGNYMYSILLACISINRYLAIVHPFTVRGLPKH-----TY 228
DB :|||:|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:
88 RGDPHWFGDLACLRFLRFVFNALHGSIILFTICISFORYLGIChPLA-----PWHRKGRRRA 143

QY	229	ALVTCGLIWAIVFLYMPLFFFIKQEYYLVQPDITTCCHDVHNTCSSSPF-----QLYYFIS	284
DB	144	AWVCVGWVLWVAQAACLPFAAATG--IQNRNVCYDL-----SPPILSTRVLPYGMA	195
QY	285	LAFPGELIPVLLIYCVAARIINLAYD-----HRWLWYVKASLLILVIFITICAPS	336
DB	196	LTVIGELLFPFALLACYCRMARLRCRODPAGPVAQERRSKAARMAVVAAVFVISFLPF	255
QY	337	NIILIITHANYNNYTDLGYFYLIALC-----LGSLNSCLDPFL-YFLMSKTRNHS	387
DB	256	HITKTAYLA---VRSPGVSCPVLETFAAAYKGTRPFASANSVLDPILFYFTQQFRQP	312
QY	388	TAYLTK	393
DB	313	HDLQK	318
RESULT 22			
interleukin-8 receptor type B - human			
C;Species: Homo sapiens (man)			
C;Date: 07-Oct-1994 #sequence revision 12-Apr-1996 #text_change 05-Nov-1999			
C;Accession: I37898, I38713, A53611, A39446			
R;Anuja, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.			
J. Biol. Chem. 269, 26381-26389, 1994			
A;Title: Comparison of the genomic organization and promoter function for human interleukin-8			
A;Reference number: I37898, MUID:95014476; PMID:7929358			
A;Accession: I37898			
A>Status: preliminary			
A:Molecule type: DNA			
A;Residues: 1-360 <RES>			
A;Cross-references: EMBL:U11869; NID:g511801; PIDN:AAB60656.1; PID:g511803			
A;Accession: I38712			
A>Status: preliminary			
A:Molecule type: mRNA			
A;Residues: 1-15 <R2>			
A;Cross-references: EMBL:U11872; NID:g511808; PIDN:AAA64380.1; PID:g511809; EMBL:U11873			
11876; NID:g511816; PID:g511817; EMBL:U11877; NID:g511818; PID:g511819; EMBL:U11878; NID:			
R;Sprengrer, H.; Lloyd, A.R.; Lautens, L.L.; Bonner, T.I.; Kelvin, D.J.			
J. Biol. Chem. 269, 11065-11072, 1994			
A;Title: Structure, genomic organization, and expression of the human interleukin-8 re			
A;Reference number: A53611; MUID:94209273; PMID:7512557			
A;Accession: A53611			
A>Status: preliminary			
A:Molecule type: DNA			
A;Residues: 6-360 <SPR>			
A;Cross-references: GB:M99412; GB:I19593			
R;Murphy, P.M.; Tiffany, H.L.			
Science 253, 1280-1283, 1991			
A;Title: Cloning of complementary DNA encoding a functional human interleukin-8 recept			
A;Reference number: A39446; MUID:91368200; PMID:1891716			
A;Accession: A39446			
A>Status: preliminary; nucleic acid sequence not shown			
A:Molecule type: mRNA			
A;Residues: 6-360 <MUR>			
A;Cross-references: GB:M73969			
C;Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8,			
C;Genetics:			
A;Gene: GDB:IL8RB, IL8RA			
A;Cross-references: GDB:127868; OMIM:146928			
A;Map position: 2q35-2q35			
C;Superfamily: vertebrate rhodopsin			
C;Keywords: G protein-coupled receptor; transmembrane protein			
Query Match 16.3%; Score 351.5; DB 2; Length 360;			
Best Local Similarity 27.5%; Pred. No. 5.9e-22;			
Matches 97; Conservative 66; Mismatches 121; Indels 69; Gaps 12;			
QY	65	NSPEEPFSALEGWTGATITVKIKCPESASHLKVNATMGYLTS-----LSKL	115
DB	9	DSPEDF-----WKG-----EDLSNTSYSTLPFLDDAAPCESUEINKYF	50
QY	116	IPAIYLLVFGVPANA-VTLMLFFRTGRICITTVFTNLAIADFLFCVTLPFKIAYHLN	174

Db 51 WIIIVLVLLSLGNSLWMLVILYVRGSRVTDVYLLNLALADLLFALTLPWAASKVN 110
Qy 175 GNNVWFEVLCRAITVIFYGNMYSILLACISINRVLAVHPTTYGRLPKHTVAL-VTC 233
Db 111 G--WIFGTFLCKVSLKEVNFYSGILLACISVDYRLAVH--ATRTLTKQRYLVKFC 166
Qy 234 GLVWATVLYMLPFFILKQEVYLVQPDITTCCHDVHNTCESSPOLYVFFSLAFEGFLIP 293
Db 167 LSIWGLSLLALPVLPRRTVYSSNVSPACVEDMGN---NTANWRMLRLILPQSGFIVP 223
Qy 294 FVLIYCYAAIIRTL----NAYDRWLWYKVASILLIIVITFCFAPSNIILII----- 342
Db 224 LLIMLCVCGFTLRLTFLKRAHMQKRAMRVIFA---VVLIFLLCWLPYNVLADTLMTQ 280
Qy 343 -----HIANVYNNTDGLFYIYIALCGLSINSLCDPLFYFLMSKTRNH 386
Db 281 VIQETCERRNHIDRALDATE-----ILGILHSLNPLIYAFIGQKFRH 323

RESULT 23

S68208
G protein-coupled receptor 12A - human
C:Species: Homo sapiens (man)
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 29-Sep-1999
C:Accession: S68208
R:An. S.; Tsai, C.; Goetzl, E.J.
FEBS Lett. 375, 121-124, 1995
A:Title: Cloning, sequencing and tissue distribution of two related G protein-coupled re
A:Reference number: S68207; MUID:96087098; PMID:7498459
A:Accession: S68208
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-365 <ANS>
A:Cross-references: EMBL:U35398; NID:g1015419; PIDN:AAA79060.1; PID:g1015419
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

Query Match 16.3%; Score 351; DB 2; Length 365;
Best Local Similarity 28.8%; Pred. No. 6.6e-22;
Matches 90; Conservative 59; Mismatches 127; Indels 36; Gaps 9;
Qy 104 MGYLT-----TKLIPAIYLLVYVGVGPANAVTLWMLFFRTRSICTT-VFY 151
Db 1 MGNITADNNTSMNCIDHTIHTQTLAPVYVTVLVVGVFPANCLISYGLYQIKARNELGYVL 60
Qy 152 TNLAIADFLFCVTLPPFKIAYHLNGNWNVFGVLCRAITVIFYGNMYSILLACISINRY 211
Db 61 CNLTAVADLFYICSLPFWLQYVQLQHDNWSHGLSCQVCGILLIYIYISVGLCCISVDRY 120
Qy 212 LAIVHPPTYRGLPKHTYALVTCGLWATVFLYMLPFFILKQEVYLVQPDITTCCHDVHNTC 271
Db 61 CNLTAVADLFYICSLPFWLQYVQLQHDNWSHGLSCQVCGILLIYIYISVGLCCISVDRY 120
Qy 212 LAIVHPPTYRGLPKHTYALVTCGLWATVFLYMLPFFILKQEVYLVQPDITTCCHDVHNTC 271
Db 121 LAVAHPPRPHQFRTLKAARVTVTVIWKELTISI-----YFLMHEEVIEDENQHRVC 172
Qy 272 ESSSPFQY----YFISLAFGLIPVLLIYCYAAIIRTL----NAYDRWLWYVKAS 322
Db 173 FEHPIQAWQAINYREL--GELFPICLLASVQGLIRAVRSHGTQSKRQIQRLV 230
Qy 323 LLILVITFCFAPSNIILIIHHANYNNNTD---GLFYIYIALCGLSINSLCDPLFYFL 379
Db 231 LSTVIVIFLACLPYHVLVLR--SWEASCDFAKGVFNAYHFSLLTTSFNCVADPVLVCF 288
Qy 380 MSKTRNHSTAYL 391
Db 289 VSETTHRDRLARL 300

RESULT 24

JC5653
G protein-coupled receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 17-Nov-1997 #sequence_revision 17-Nov-1997 #text_change 21-Jul-2000
C:Accession: JC5653

R:Ferrer-Martinez, A.; Felipe, A.; Mata, J.F.; Casado, F.J.; Pastor-Anglada, M.
Biochem. Biophys. Res. Commun. 238, 107-112, 1997
A:Title: Molecular cloning of a bovine renal G-protein coupled receptor gene (BRGR): R
A:Reference number: JC5653; MUID:97445134; PMID:9299461
A:Accession: JC5653
A:Molecule type: mRNA
A:Residues: 1-361 <PER>
A:Cross-references: GB:U88366; NID:92827875; PIDN:AAC05611.1; PID:g2827876
A:Experimental source: renal epithelial cell
C:Superfamily: G protein-coupled receptor 4
C:Keywords: glycoprotein; phosphoprotein
F:24-46/Domain: transmembrane #status predicted <TM1>
F:58-79/Domain: transmembrane #status predicted <TM2>
F:95-117/Domain: transmembrane #status predicted <TM3>
F:137-159/Domain: transmembrane #status predicted <TM4>
F:189-207/Domain: transmembrane #status predicted <TM5>
F:229-253/Domain: transmembrane #status predicted <TM6>
F:265-289/Domain: transmembrane #status predicted <TM7>
F:3,62/Binding site: carboxylate (Asn) (covalent) #status predicted
F:134,218,293,313/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status
F:204,221,328/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pr

Query Match 16.1%; Score 347; DB 2; Length 361;
Best Local Similarity 27.4%; Pred. No. 1.4e-21;
Matches 86; Conservative 61; Mismatches 127; Indels 40; Gaps 8;
Qy 104 MGYLTS-----SLSTKLIPAIYLLVYVGVGPANAVTLWMLFFRTRSICTT-VFY 151
Db 1 MGNITADNNTSMNCIDHTIHTQTLAPVYVTVLVVGVFPANCLISYGLYQIKARNELGYVL 60
Qy 152 TNLAIADFLFCVTLPPFKIAYHLNGNWNVFGVLCRAITVIFYGNMYSILLACISINRY 211
Db 61 CNLTAVADLFYICSLPFWLQYVQLQHDNWSHGLSCQVCGILLIYIYISVGLCCISIDRY 120
Qy 212 LAIVHPPTYRGLPKHTYALVTCGLWATVFLYMLPFFILKQEVYLVQPDITTCCHDVHNTC 271
Db 121 LAVAHPPRPHQFRTLKAAMGVSAIIVWKELTISI-----YFLMHEEVVEDADRHRVC 172
Qy 272 ESSSPFO-----LYFISLAFGLIPVLLIYCYAAIIRTL----NAYDRWLWYVK 320
Db 173 FEHPIEPRGNGINYRFL---VGFLFPICLLASVQGLIRAVRSHGTQSKRQIQIOR 228
Qy 321 ASLLILVITFCFAPSNIILIIHHANYNNNTD---GLFYIYIALCGLSINSLCDPLFY 377
Db 229 LVISTVIVIFLACLPYHVLVLR--SLWESSCDFAKGVFNAYHFSLLTTSFNCVADPVLV 286
Qy 378 FLMSKTRNHSTAYL 391
Db 287 CFVSETTHRDRLARL 300

RESULT 25

S63666
platelet activating factor receptor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S63666
R:Ishii, S.; Matsuda, Y.; Nakamura, M.; Waga, I.; Kume, K.; Izumi, T.; Shimizu, T.
Biochem. J. 314, 671-678, 1996
A:Title: A murine platelet-activating factor receptor gene: cloning, chromosomal local
A:Reference number: S63666; MUID:96239129; PMID:8670084
A:Accession: S63666
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-341 <ISH>
A:Cross-references: EMBL:D50872; NID:g1256924; PIDN:BAA09468.1; PID:g1256925
C:Superfamily: ATP receptor P2u

Query Match 16.0%; Score 346; DB 2; Length 341;
Best Local Similarity 29.1%; Pred. No. 1.6e-21;
Matches 92; Conservative 49; Mismatches 141; Indels 34; Gaps 10;
Qy 115 LIPAIYLLVYVGVGPANAVTLWM---LFFRTRSICTTVFVTNLAIADFLFCVTLPPFKIAY 171

Db 17 LFPIVYSVIFILGVANGYVWLVFANLYPSKXKNEIKIFMVNLMTADLLFLITLPLMIVY 76
Qy 172 HLGNNWVFGVLCRATTVIFGYNNMYSILLACISINRYLAIVHPFTYRGLPKHTVALV 231
Db 77 YNNEGDWLLPFLCNVACGLFINTYCSVAFGLVITNRYQAVAYPIKTAQATTKRGIS 136
Qy 232 TCGLVWATVLYMLPFFILKOEYVILV-----OPDITTCDDVHNTCESSPPOLYFISLA 286
Db 137 LSLIIWISI-VATASYFLATDSTNLVFNKOGSGNITRCFE-HYEPYSVPILVWHVFIATC 194
Qy 287 FFGFLIPPLVLIYCAAIITRL-----NAYDRHLMWYKASLLILVIFTICFAPS 336
Db 195 FF---LVFPLIFYNLVIIHTLLQPMRQQRKAGVKRRALMW---CTVLAVFIICFVPH 248
Qy 337 NIILI---IHANYNNNTDGLFYLIALLCGLSNCLDPPFLY-FLMSKTRNHSYALT 392
Db 249 HVQVLPTWLAELGVQTNPHQAINDAHQITLCLLSTNCLDVPVIYCFLLTKKFKH----LS 304
Qy 393 KXKNDLREQQPSORT 408
Db 305 EKFSMRSSRKRCSRAT 320

RESULT 26

A48921
Interleukin-8 receptor type B - mouse
N:Alternate names: G-protein coupled receptor Gpcrl6
C:Species: Mus musculus (house mouse)
C>Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text change 05-Nov-1999
R:Accession: A48921; A53677; I49348; I55421; H48909; I53774
R:Cerretti, D.P.; Nelson, N.; Kozlosky, C.J.; Morrissey, P.J.; Copeland, N.G.; Gilbert, Genomics 18, 410-413, 1993
A:Title: The murine homologue of the human interleukin-8 receptor type B maps near the i
A:Reference number: A48921; MUID:94117014; PMID:8288247
A:Accession: A48921
A:Molecule type: DNA
A:Residues: 1-359 <CER>
A:Cross-references: GB:L23637; NID:G435093; PIDN:AAA39305.1; PID:G435094
R:Suzuki, H.; Prado, G.N.; Wilkinson, N.; Navarro, J.
J. Biol. Chem. 269, 18263-18266, 1994
A:Title: The N terminus of interleukin-8 (IL-8) receptor confers high affinity binding t
A:Reference number: A53677; MUID:94308043; PMID:7518426
A:Accession: A53677
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-359 <SUZ>
A:Cross-references: GB:L26549
A:Note: sequence extracted from NCBI backbone (NCBIP:149812)
R:Lee, J.; Cacalano, G.; Camarato, T.; Toy, K.; Moore, M.W.; Wood, W.I.
J. Immunol. 155, 2158-2164, 1995
A:Title: Chemokine binding and activities mediated by the mouse IL-8 receptor.
A:Reference number: 149348; MUID:95363183; PMID:7636264
A:Accession: 149348
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <RES>
A:Cross-references: EMBL:U31207; NID:G950174; PIDN:AAC52239.1; PID:G950175
R:Bozic, C.R.; Gerard, N.P.; von Uexkull-Guldenband, C.; Kolakowski, L.F.
J. Biol. Chem. 269, 29355-29358, 1994
A:Title: The murine interleukin 8 type B receptor homologue and its ligands. Expression
A:Reference number: 155421; MUID:95050766; PMID:7961909
A:Accession: 155421
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-359 <RE2>
A:Cross-references: GB:Li3239; NID:G293665; PIDN:AAA62109.1; PID:G293666
R:Wilkie, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland, N.C
Genomics 18, 175-184, 1993
A:Title: Identification, chromosomal location, and genome organization of mammalian G-p
A:Reference number: A48909; MUID:94116980; PMID:8288218
A:Accession: A48909
A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 145-258 <WIL>
A:Cross-references: GB:L20337; NID:G438800; PIDN:AAA16853.1; PID:G438801
R:Harada, A.; Kuno, K.; Nomura, H.; Mukaida, N.; Murakami, S.; Matsushima, K.
Gene 142, 297-300, 1994
A:Title: Cloning of a cDNA encoding a mouse homolog of the interleukin-8 receptor.
A:Reference number: 153774; MUID:94252584; PMID:8194768
A:Accession: 153774
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-359 <RE3>
A:Cross-references: GB:D17630; NID:G493671; PIDN:BAA04536.1; PID:G493672
C:Genetics:
A:Gene: IL8rb
A:Introns: #status absent
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:49-74/Domain: transmembrane #status predicted <TM1>
F:84-106/Domain: transmembrane #status predicted <TM2>
F:120-141/Domain: transmembrane #status predicted <TM3>
F:163-182/Domain: transmembrane #status predicted <TM4>
F:213-234/Domain: transmembrane #status predicted <TM5>
F:251-271/Domain: transmembrane #status predicted <TM6>
F:308-328/Domain: transmembrane #status predicted <TM7>

Query Match 16.0%; Score 345.5; DB 2; Length 359;
Best Local Similarity 29.8%; Pred. No. 1.9e-21;
Matches 96; Conservative 59; Mismatches 120; Indels 47; Gaps 11;
Qy 111 LSTKLPDAIYLLVFFVGVGVPANAVTLMLPFRTRISIC--TTVFYTNLAIAIDFLFCVTLPPFK 168
Db 45 INSAVVVIVYVTLTSLSLVGNLMVLVILY-NRSTCSVTVDYLLNLNLAIDFLFAFLTLPVW 103
Qy 169 IAYHLGNWVFGVLCRATTVIFGYNNMYSILLACISINRYLAIVHPFTYRGLPKHTY 228
Db 104 AASKVNG--WTFGSTLCKIFSYKVEVTFYSSVLLACISMDRYLAIVHATSTLIQKHLV 161
Qy 229 ALVTCGLVWATVLYMLPFFILKOEYVILVQPDITTCDDVHNTCESSPPOLYFISLAFF 288
Db 162 KKV-CIAMWLLSVILALPILLRNPKVNLSTLCYEDVGN---NTSLRVLRLPQT 217
Qy 289 GLFIPFLIYCAAIITRL-----NAYDRHLMWYKASLLILVIFTICFAPSNIILIH 344
Db 218 GLFVPLIIMLCYCFGLTTLTKAHMGOKHRMVRIFA---VVLVFLCLPLYNLVLF-- 271
Qy 345 ANYYNNTCGLYFIYLI-----AL-----CLGSLNSCLDPLFLMSKTRNHS- 387
Db 272 -----TDTLMRTKLKETCERRDDIDKALNATEILGFUHSCLNPIIYAFIQGKFRHL 324
Qy 388 ---TAYLTXXXNDLREQGPS 405
Db 325 LKIMATYGLVSKBFLAKEGRPS 346

RESULT 27

B57641
G protein-coupled receptor 4 - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 08-Feb-1996 #sequence revision 08-Feb-1996 #text change 29-Sep-1999
C:Accession: B57641
R:Mahadevan, M.S.; Baird, S.; Bailly, J.E.; Shutler, G.G.; Sabourin, L.A.; Tsilfidis, C.
Genomics 30, 84-88, 1995
A:Title: Isolation of a novel G protein-coupled receptor (GPR4) localized to chromosome
A:Reference number: A57641; MUID:96129306; PMID:8595909
A:Accession: B57641
A:Status: preliminary; nucleic acid sequence not shown; translation not shown; signific
A:Molecule type: DNA
A:Residues: 1-362 <MAH>
A:Cross-references: GB:U22108; NID:G722282; PIDN:AAA98458.1; PID:G722283
C:Genetics:
A:Introns: #status absent
C:Superfamily: G protein-coupled receptor 4
C:Keywords: G protein-coupled receptor

C;Accession: S26667
R;Dobner, T.; Wolf, I.; Emrich, T.; Lipp, M.
Eur. J. Immunol. 22, 2795-2799, 1992
A;Title: Differentiation-specific expression of a novel G protein-coupled receptor
A;Reference number: S26667; MUID:93049615; PMID:1425907
A;Accession: S26667
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-372 <DOB>
A;Cross-references: EMBL:X68149; NID:g29459; PIDN:CAA48252.1; PID:g29460
C;Genetics:
A;Gene: GDB:BLR1
A;Cross-references: GDB:136235; OMIM:601613
A;Map position: 15q26.1-15q26.1
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.9%; Score 344; DB 2; Length 372;
Best Local Similarity 33.0%; Pred. No. 2.6e-21;
Matches 104; Conservative 51; Mismatches 130; Indels 30; Gaps 15

Qy 98 HVKNATMGYLTSSLSKLIPIAYLLFVGVGPANAVITLWML--FPRTSICITVFVTNLA 155
Db 36 HLCPEATGGLPMASFKAIVPVAYSLIFLLGVIGNVLVILIERHQTRS-STETFLFLHA 94
Qy 156 IADFLFCVTLPFKIAYHLNGNNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIV 215
Db 95 VADLLLVFLPFAVAGSVG--VWLGTFLCKTVIALHKVNFYCSSLLACIADRYLAIV 152
Qy 216 HP-FTYRGLPKHTVAL---VTCGLVWATVFLYMLP---FFILKQBYVLVQDPDITTC-HDV 267
Db 153 HAVHAYR---HRRLLSIHITCGTIVLWGFLLALPEILFAKVSQGH--NNSLPRCTTFSQ 206
Qy 268 HNTCESSSPQLXYFISIAFPGLPIPFVLLIYCYAAIIRTNAYDHR--WLMYVKASLLI 325
Db 207 ENQETHAWFTSRFLYHVA--GFLPDLVWGVCVGVVHRUQAORRRQKAVRAVLV 264
Qy 326 LVFIITTCFAPSNIILIIHANY--YNNTDGL--YFIYLIALC--LGSLSNCLDPPFLY-F 378
Db 265 TSIFFLCWSPVHIVIFDLTLARLKAVDNTCKNGSLPVAITWCEFLGLAHCCLNPMFLYT 324
Qy 379 LMSKTRNHSTAYLTK 393
Db 325 AGVKFRSDLSRLTK 339

RESULT 30
151372
angiotensin II receptor - turkey
C;Species: Meleagris gallopavo (common turkey)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 24-Nov-1999
C;Accession: 151372; PN0449
R;Murphy, T.J.; Nakamura, Y.; Takeuchi, K.; Alexander, R.W.
Mol. Pharmacol. 44, 1-7, 1993
A;Title: A cloned angiotensin receptor isoform from the turkey adrenal gland
A;Reference number: 151372; MUID:93341466; PMID:8341266
A;Accession: 151372
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-359 <GBR>
A;Cross-references: GB:L23203; NID:g349735; PIDN:AAA03560.1; PID:g349736
R;Carsia, R.V.; McIlroy, P.J.; Kowalski, K.I.; Tilly, J.L.
Biochem. Biophys. Res. Commun. 191, 1073-1080, 1993
A;Title: Isolation of turkey adrenocortical cell angiotensin II (AII) receptor
A;Reference number: PN0449; MUID:93221469; PMID:7916599
A;Accession: PN0449
A;Molecule type: mRNA
A;Residues: 53-91, 'C', 93-162, 'SSFIVY', 171-291 <CAR>
A;Cross-references: GB:858041; NID:g2939373; PIDN:AAB26041.1; PID:g299374
A;Experimental source: liver
C;Superfamily: vertebrate rhodopsin
C;Keywords: glycoprotein; phosphoprotein; transmembrane protein
F;53-64/Domain: intracellular #status predicted <INI>

A:Molecule type: mRNA
A:Residues: 1-342 <NAK>
A:Cross-references: GB:D10202; GB:D90433; NID:G219975; PIDN:BAA01050.1; PID:G219976
R:Sugimoto, T.; Tsuchimoto, H.; McGregor, C.G.A.; Mutoh, H.; Shimizu, T.; Kurachi, Y.
Biochem. Biophys. Res. Commun. 189, 617-624, 1992
A:Title: Molecular cloning and characterization of the platelet-activating factor receptor
A:Reference number: JCI359; MUID:93112021; PMID:1281995
A:Accession: JCI359
A:Molecule type: mRNA
A:Residues: 1-315, 'N', 317-342 <SUG>
A:Experimental source: heart
A:Note: the authors translated the codon AAT for residue 316 as Lys
R:Seffried, C.E.; Schweickart, V.L.; Godiska, R.; Gray, P.W.
Genomics 13, 832-834, 1992
A:Title: The human platelet-activating factor receptor gene (PTAFR) contains no introns
A:Reference number: A42831; MUID:92347886; PMID:1322356
A:Accession: A42831
A:Molecule type: DNA
A:Residues: 1-226, 'TG', 229-342 <SE>
A:Cross-references: GB:W88177; NID:G190697; PIDN:AAA60214.1; PID:G190698
A:Note: sequence extracted from NCBI backbone (NCBIN:109813, NCBI:P:109814)
R:Chase, P.B.; Halonen, M.; Regan, J.W.
Am. J. Respir. Cell Mol. Biol. 8, 240-244, 1993
A:Title: Cloning of a human platelet-activating factor receptor gene: evidence for an
A:Reference number: I51923; MUID:93192035; PMID:8383507
A:Accession: I51923
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-342 <RES>
A:Cross-references: GB:S56396; NID:G298580; PIDN:AB25755.1; PID:G298581
C:Genetics:
A:Gene: GDB:PTAFR
A:Cross-references: GDB:128806; OMIM:173393
A:Map position: lp35-1p34.3
C:Superfamily: ATP receptor p2u
C:Keywords: G protein-coupled receptor; transmembrane protein
F:17-38/Domain: transmembrane #status predicted <TRI>
F:54-75/Domain: transmembrane #status predicted <TII>
F:92-113/Domain: transmembrane #status predicted <TII>
F:134-155/Domain: transmembrane #status predicted <TIV>
F:184-205/Domain: transmembrane #status predicted <TVI>
F:233-253/Domain: transmembrane #status predicted <TVI>
F:277-297/Domain: transmembrane #status predicted <VII>

Query Match 15.9%; Score 342.5; DB 2; Length 342;
Best Local Similarity 28.3%; Pred. No. 3.2e-21;
Matches 96; Conservative 50; Mismatches 132; Indels 61; Gaps 14;

Qy 106 YLTSSLSTKLIPAIYLLVVGVPANAVTLMLPFRTRISIC-----TTVFYNLAIADEL 160
Db 8 HMDSEPRYTLFPIVYSIIFVLGIANGYVLWV--FARLYPCKKFEIKIFVNLTMADML 65
Qy 161 FCVTLPKIAYHLNGNNVGFVEVLCTATTVFYGNWYCSILLACISINRYLAIIVHPFTY 220
Db 66 FLITPLMIWVYQNGNWLKPLCNVAGCLUFFINTYCSVAFGLVITVYNNRFOAVTRPIKT 125
Qy 221 RGLPKHTVALVTCGLWATVFLYMLPFILKQEYVLV-----QPD-----ITTCHDVH 268
Db 126 AQANTRKRGISLSLVIWAI-----VGAASFLILDSTNTVPDSAGSNVTRCFE-H 176
Qy 269 NTCSSSPFQLYPFISIAFFGLPIFVLIIYCYAAIIRTL-----NA-YDHRWLWY 318
Db 177 YEGSVFPLIIHIFI--VFSEFLVFLIILFCNLVIIRTLMLQPVQOQORNAEVKRRALWM 233
Qy 319 VKASLLIIVFTICFAPSNI-----LIHHANYNNYNDGLFYFLIALCLGLSLNS 370
Db 234 V--CTVLAVFIICFVPHVHVLQPLWTLAEGLFQDSKFHQAINDA-----HQVTLCLLSTNC 286
Qy 371 CLDPFLY-FLMSKTRNHSTAYLTXXNDLREQGQPSORT 408
Db 287 VLDPVIVCYFLTKKPRKH----LTEKFYSMRSSRCKSRAT 321

Eur. J. Biochem. 221, 211-218, 1994
 A;Title: Cloning, expression and tissue distribution of rat platelet-activating-factor-
 A;Reference number: S43252; MUID:94222063; PMID:8168510
 A;Accession: S43252
 A;Molecule type: mRNA
 A;Residues: 1-341 <BIT>
 A;Cross-references: GB:U04740; NID:G470384; PIDN:AAA18422.1; PID:G470385
 C;Superfamily: ATP receptor P2u

Query Match	15.8%	Score 340;	DB 2;	Length 341;
Best Local Similarity	30.1%	Pred. No. 5.1e-21;		
Matches	97;	Conservative 47;	Mismatches 132;	Indels 46; Gaps 12;

Qy	115	LIPAIYLLVFVGVGPANAVTLWM---	LPFRTRSICTTVFYTNLAIADELFVCTLPFKIAY	171
Db	17	LFPIVYSVIFVLGVVANGVWLWFATLYPSKKLNEIKIFMVNLTIVADLLFLMTLPWLIVY		76
Qy	172	HLGNWVFEGLCRATTVFYGNMYCSILLACISINRYLAIVHPF-----	TYRGLP	224
Db	77	YSNEGDWIVHKFLCNLACGLFFINTYCSVAFGLGVITYNRYQAVYPIKTAQATTKRGI-		135
Qy	225	KHYVALVTCGLVWATVFLYMLPFFILKOEYLVOPD----	ITTCDDVHNTCESSSPFOLY	280
Db	136	--TSLSV---IWISIAATASYFLATDSTNVVPKKGSGNITRCFE-HYEPYSVPILVVH		188
Qy	281	YFISLAFPGFLIPFVLIYCYAAIIRTL-----	NAYDHRMLWYVKASLLILVIFT	330
Db	189	IFITSCFF--LVFELIFYCNMVIHTLTPVRQQRKPEVKRRALMWV---CTVLAVFV		242
Qy	331	ICFAPSNIILI---IHHANYNNYNDGLFYIYLIACLGSLNSCLDPLLY-FLMSKTRNH		386
Db	243	ICFVPHVHVQLPWTLAELGYQTNFHQAINDAHQITCLLSTNCVLDVPVIYCFLTAKFKKH		302
Qy	387	STAVLTXXXNDLREGQPSORT		408
Db	303	----LSEKFYSMRSSRKCSRAT		320

Search completed: June 24, 2003, 12:06:37
 Job time : 29.0319 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2003, 12:05:46 ; Search time 32.5399 Seconds
(without alignments)
1356.747 Million cell updates/sec

Title: US-09-208-629F-6
Perfect score: 2157
Sequence: 1 CSMLQISXRLRDGTQVIKM.....AYLTXXNLDLRGQPSQRT 408

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523	24.2	425	10	US-09-782-980-80
2	523	24.2	425	10	US-09-884-430-4
3	471.5	21.9	359	10	US-09-943-718-6
4	460	21.3	359	9	US-10-190-469-1
5	460	21.3	359	9	US-10-094-417-10
6	460	21.3	359	9	US-09-782-974C-76
7	460	21.3	359	10	US-09-739-151-2
8	428	19.8	537	9	US-10-311-956-4
9	426	19.7	362	9	US-10-092-135-4
10	424	19.7	362	9	US-10-092-135-3
11	424	19.7	362	9	US-09-779-679-28
12	413.5	19.2	339	9	US-10-251-385-182
13	410.5	19.0	339	9	US-09-828-478-4
14	410.5	19.0	339	9	US-10-251-385-32
15	410.5	19.0	339	10	US-09-848-889-12
16	410.5	19.0	339	10	US-09-788-133-2
17	410.5	19.0	361	9	US-10-222-024-2
18	410.5	19.0	361	9	US-10-251-385-78
19	410.5	19.0	361	9	US-10-251-385-206

20	410.5	19.0	367	9	US-09-828-478-6	Sequence 6, Appli
21	409.5	19.0	373	9	US-10-092-135-7	Sequence 7, Appli
22	405.5	18.8	348	10	US-09-827-937A-17	Sequence 17, Appli
23	399	18.5	373	9	US-10-092-135-6	Sequence 6, Appli
24	398	18.5	373	9	US-10-092-135-5	Sequence 5, Appli
25	397	18.4	302	9	US-10-024-494-30	Sequence 30, Appli
26	396.5	18.4	299	9	US-10-270-144-4	Sequence 4, Appli
27	395	18.3	366	9	US-09-779-679-25	Sequence 25, Appli
28	384.5	17.8	360	10	US-09-943-718-4	Sequence 4, Appli
29	382	17.7	365	9	US-09-077-173A-2	Sequence 2, Appli
30	376.5	17.5	346	9	US-09-828-478-2	Sequence 2, Appli
31	376.5	17.5	346	9	US-09-779-679-2	Sequence 2, Appli
32	376.5	17.5	346	9	US-09-779-679-26	Sequence 26, Appli
33	376.5	17.5	346	10	US-09-826-791-6	Sequence 6, Appli
34	376.5	17.5	346	10	US-09-866-230-7	Sequence 7, Appli
35	376.5	17.5	346	10	US-09-866-230-9	Sequence 9, Appli
36	375	17.4	344	9	US-10-024-494-8	Sequence 8, Appli
37	374.5	17.4	330	10	US-09-826-791-2	Sequence 2, Appli
38	372.5	17.3	337	9	US-09-828-478-5	Sequence 5, Appli
39	372.5	17.3	337	10	US-09-866-230-8	Sequence 8, Appli
40	368.5	17.1	358	10	US-09-826-508-10	Sequence 10, Appli
41	365	16.9	337	9	US-09-779-679-27	Sequence 27, Appli
42	364.5	16.9	337	9	US-10-167-192-3	Sequence 3, Appli
43	364.5	16.9	374	9	US-10-242-499-3	Sequence 3, Appli
44	361	16.7	341	9	US-10-270-587-3	Sequence 3, Appli
45	361	16.7	358	9	US-09-974-298-86	Sequence 86, Appli

ALIGNMENTS

RESULT 1

US-09-782-980-80
; Sequence 80, Application US/09782980
; Patent No. US20020072089A1
; GENERAL INFORMATION:
; APPLICANT: Khodadoust, Mehran M.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Busfield, Samantha J.
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Gu, Wei
; APPLICANT: White, David
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIFE, TRASH, BDSF, LRSG, AND
; TITLE OF INVENTION: SMST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-121CP
; CURRENT APPLICATION NUMBER: US/09/782,980
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: PCT/US00/02125
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: 09/448,076
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: 09/276,400
; PRIOR FILING DATE: 1999-03-25
; PRIOR APPLICATION NUMBER: 60/117,580
; PRIOR FILING DATE: 1999-01-27
; PRIOR APPLICATION NUMBER: 09/014,195
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/014,348
; PRIOR FILING DATE: 1998-01-27
; PRIOR APPLICATION NUMBER: 09/086,892
; PRIOR FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 09/296,208
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 09/063,950
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 09/561,381
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/561,810
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 09/087,121


```

; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-943-718-6

Query Match      21.9%; Score 471.5; DB 10; Length 359;
Best Local Similarity 36.6%; Pred. No. 7.7e-31;
Matches 102; Conservative 47; Mismatches 111; Indels 19; Gaps 4;

QY 112 STKLPAIYLLVFVGVGPANAVTLWMLFFTRISCTTTFVTNLAIADFLFCVTLPKIAY 171
DB 11 SPQVPAIYGLVAVGLPANGLAWLAVTRPLPSTILLMNLAVALDLALVLPRLAY 110
QY 172 HLGNNMVFGEVLCRATTVIFYGNMYSIILLACISINRYLAIVHPPTYRGLPKHTYALV 231
DB 111 HLRGQRPFGCEACRVATAALYGHMYSVLLAAVSLDRYLAIVHPLRABARGLRQTTC 170
QY 232 TCGLVWATVFLYMLPFFILKQEYVLVQPDITTCVDVHNTC-----ESSPPFQIYYP 282
DB 171 LCLVAMLSAATLALPLTLHRQNFLLAP-IACC--VMRCPWLSRTPGTGERPSSAMLSWA 227
QY 283 ISLAFTGFLIPFVLIYCYAAIIRTLNAYDHRWLVYKASULLIIVFTTICFAPSNIILII 342
DB 228 ASL-----PLLAMGLCYGTTLRALAANGORYSHALRTLALVLFSAVASFTPSNVLVL 280
QY 343 HHANYNNYNTDGLYFIYLIACLSLNSCLDPFLYFLMS 381
DB 281 HYNPSPGANGNLYGAVPSLALSTLNSCLDPPHYIYVVS 319

RESULT 4
US-10-190-469-1
; Sequence 1, Application US/10190469
; Publication No. US20030013155A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 14400 Receptor, A No. US20030013155A1e1 G-Protein Coupled Receptor
; FILE REFERENCE: 5800-7, 035800/169196
; CURRENT APPLICATION NUMBER: US/10/190,469
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: US/09/137,063A
; PRIOR FILING DATE: 1999-08-20
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-190-469-1

Query Match      21.3%; Score 460; DB 9; Length 359;
Best Local Similarity 34.4%; Pred. No. 6.7e-30;
Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;

QY 101 NATMGYLTSSLTKLIPAIYLLVFVGVGPANAVTLWMLFFR--TRSICTTVFVTNLAIAD 158
DB 11 NATLQMLRNPALPAVALPVVYSLVAASIPGNLPSLWLCRRMGPRS-PSVIFMINLSVTD 69
QY 159 FLFCVTLPKIAYHLNGNNMVFGEVLCRATTVIFYGNMYSIILLACISINRYLAIVHPF 218
DB 70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVTVAFYANNMYSILLTMTCSIVERFLGVLYPL 129
QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCVDVHNT-TCSESSSPF 277
DB 130 SSKWRERRRYAVAACAGTWLTLTALSPLARTDLYTPVHALGIITCFDVLKWTMLPSVAM 189
QY 278 QLYYFISLAFFGLIPFVLIYCYAA----IIRTLNAYD-HRWLVYKASULLIIVFTTIC 332
DB 190 WAVFLFTIFILLFPIPVITVACYTATILKLLRTEAHGREQRRRAVGLAAVVLLAFVTC 249
QY 333 FAPSNIIILIIH-HANYNNYNTDGLYFIYLIACLSLNSCLDPFLYFLMSK 382
DB 250 FAPNNEVLLAHIVSRIFYGKS--YHVVYKLTCLSLNCLNCLDPFFVYFASR 298

RESULT 6
US-09-782-974C-76
; Sequence 76, Application US/09782974C
; Publication No. US20030082534A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; APPLICANT: Lind, Peter
; APPLICANT: Wood, Linda S.
; APPLICANT: Parodi, Luis A.
; TITLE OF INVENTION: No. US20030082534A1e1 G Protein Coupled Receptor
; FILE REFERENCE: 411USPHRM311
; CURRENT APPLICATION NUMBER: US/09/782,974C
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/165,838
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 09/714,449
```

```

; LENGTH: 359 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-094-417-10

Query Match      21.3%; Score 460; DB 9; Length 359;
Best Local Similarity 34.4%; Pred. No. 6.7e-30;
Matches 100; Conservative 54; Mismatches 125; Indels 12; Gaps 7;

QY 101 NATMGYLTSSLTKLIPAIYLLVFVGVGPANAVTLWMLFFR--TRSICTTVFVTNLAIAD 158
DB 11 NATLQMLRNPALPAVALPVVYSLVAASIPGNLPSLWLCRRMGPRS-PSVIFMINLSVTD 69
QY 159 FLFCVTLPKIAYHLNGNNMVFGEVLCRATTVIFYGNMYSIILLACISINRYLAIVHPF 218
DB 70 LMLASVLPFQIYYHCNRHHWVFGVLLCNVTVAFYANNMYSILLTMTCSIVERFLGVLYPL 129
QY 219 TYRGLPKHTYALVTCGLVWATVFLYMLPFFILKQEYVLVQPDITTCVDVHNT-TCSESSSPF 277
DB 130 SSKWRERRRYAVAACAGTWLTLTALSPLARTDLYTPVHALGIITCFDVLKWTMLPSVAM 189
QY 278 QLYYFISLAFFGLIPFVLIYCYAA----IIRTLNAYD-HRWLVYKASULLIIVFTTIC 332
DB 190 WAVFLFTIFILLFPIPVITVACYTATILKLLRTEAHGREQRRRAVGLAAVVLLAFVTC 249
QY 333 FAPSNIIILIIH-HANYNNYNTDGLYFIYLIACLSLNSCLDPFLYFLMSK 382
DB 250 FAPNNEVLLAHIVSRIFYGKS--YHVVYKLTCLSLNCLNCLDPFFVYFASR 298

RESULT 5
US-10-094-417-10
; Sequence 10, Application US/10094417
; Publication No. US20030045685A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030045685A1e1 Receptors
; FILE REFERENCE: 018781-008110US
; CURRENT APPLICATION NUMBER: US/10/094,417
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/802,803
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 60/276,649
; PRIOR FILING DATE: 2001-03-16
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 10
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR341
US-10-094-417-10
```

[illegible]

Query Match	19.8%;	Score 428;	DB 9;	Length 537;
Best Local Similarity	36.0%;	Pred. No. 4.2e-27;		
Matches	105;	Conservative 41;	Mismatches 102;	Indels 44; Gaps 12
Qy	115	LIPAIYLLVFVGVPANAVILMLLFFRTSR--ICTVTFVYNLAIDFLPCVTLPPKFIAYHL	173	
Db	45	LLPVSYSAVFMVGPLNIAAMWIFIAKRPWNPTVYMFENALSDTLLVLSLPTLLVYYIA	104	
Qy	174	NGNNWFGEVLCRATTWIFVGNMYCISILLACISINRYLAIVHPFT--YRGLPKHTYALV	231	
Db	105	DKNWNPFGEVLCKLVRFLFYANLYSSILFLTCISVHRVYGVCHPTISLRMNKHAHY--V	162	
Qy	232	TCGLVWATVFLYMLPPFILKQEYVLYVPDI--TTCHDVHNTCESSPPQLYFYSLAFPG	289	
Db	163	ICALVWLSVTLCLEPNLI---FVTVSPKVKNTTCHDT---TRPEDFARYVEYSTALMC	214	
Qy	290	FL--IPFVLLIYCVAAILR-----TLNAYDHRWLMWVKVKSLLILVIFTCFAP	335	
Db	215	LLFGICPLITAGCGLMTRMLKPIVSGNQOTLPSYKR---SKTITFVMAIFALCMP	271	
Qy	336	SNIIILIIHHANYNNYNTDGL--YFI-----YLIALCLGSLNSCLDPELYFL	379	

PRIOR APPLICATION NUMBER: USSN 60/181157
; PRIOR FILING DATE: 2000-02-09
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-779-679-28

Query Match 19.7%; Score 424; DB 9; Length 362;
Best Local Similarity 28.9%; Pred. No. 5.8e-27;
Matches 101; Conservative 59; Mismatches 143; Indels 46; Gaps 9;
QY 77 GWGATITVKIKPEESASHLVKQATMGVLTSSLSKLPALYLLVFFVGVGPANAVTLW 136
DB 21 GWAAGNATTKCSITKTFQFY-----LPTVYILVFIITGFLGNSVAIW 63
QY 137 MLFFRTRISIC--TVFYTNLAIAADFLFCVTLTPFKIAVHLNGNNWVFGVLCRAITTVIFYGN 195
DB 64 MFVFMHPWGSISVYFNELALADFLVLTLPALIFYFNKTDWIFGDVMCKLQRFIFHVN 123
QY 196 MYCSILLACISINRYLAIVHPPTYRGLPKHTVALVTCGLVWATVFLYMLPFFILKOEYY 255
DB 124 LYGSIPLTICISVHRVTGVVHPLKSLGRKKQNAVYVSSLVWALVAVIAPIL-----FY 178
QY 256 ----LVQPDITTDCHDVHNTCESSPPOLYYFISLAFEGFLIPVLLIYCYAAIIRTLNAY 311
DB 179 SGTVGRNKITCYD--TTADEYLSYFVYVSMCTTVFMCFIPVILGCGGLIVKALIVY 236
QY 312 D-----HRLMVYKASLLIVFTICFAPSNIILIIH-HANYNN-----TDGLVF 357
DB 237 DLDSPLRRKSIYV---IIVLTVFVSYLPFHVMTKLNRLRDLQTPQWCAFNKVYA 293
QY 358 IYIALCLGSLNSCLDPFLFLMSKTRNHSTAYLTXXNDLREQGPSQ 406
DB 294 TYQVTRGLASLNSCLDPILYFLAGDTFRRRLSRATRKSSRRSEPNVQSK 342

RESULT 12
US-10-251-385-182
; Sequence 182, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 182
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-182

Query Match 19.2%; Score 413.5; DB 9; Length 339;
Best Local Similarity 31.0%; Pred. No. 3.9e-26;
Matches 108; Conservative 49; Mismatches 130; Indels 61; Gaps 13;
QY 62 APNSFEFPFSALEGTGATITVKIKPEESASHLVKQATMGVLTSSLSKLPALYLL 121
DB 7 APGLITNFSLATAE-----QCQGE-----TPLENMFLASFYL 39
QY 122 LVFVVGVPANAVTLWMLFFRTRISICT--TVFYTNLAIAADFLFCVTLTPFKIAVHLNGNNW 179

DB 40 LDFILALVGNITLALW-LFIRDHKSQTPANVFLMHLAVADLSCLVLPTRLVYVHFSGNHWP 98
QY 180 FGEVLCRAITTVIFYGNMYCSILLACISINRYLAIVHPPTYRGLPKHTVALVTCGLVWAT 239
DB 99 FGEIACRLTGFLFYLNWYASIIYFLTCISADRFALIVHPVKSUKLRPLVAHLACAFWV 158
QY 240 VFVYMLPFFILKOEYYLVQPDITTDCHDVHNTCESSPPOLY-----YFISLAFEGFLI 292
DB 159 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203
QY 293 PFVLLIYCYAAIIRTLN---AYDHRWLWYVKASL-LILVIFTICFAPSNI---ILIIHA 345
DB 204 PTTTTCVLLIIRSLRQGLRVEKRLKTKAKRMIAIVLAIFLVCFVPYVHNSRVYVHLHYR 263
QY 346 NYYNNYNTDG--LYFYIYIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
DB 264 SHGASCATORILALANRITSCITSLNGALDPIIMYFVFAEKFRHALCNL 311

RESULT 13
US-09-828-478-4
; Sequence 4, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-4

Query Match 19.0%; Score 410.5; DB 9; Length 339;
Best Local Similarity 30.7%; Pred. No. 6.8e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;
QY 62 APNSFEFPFSALEGTGATITVKIKPEESASHLVKQATMGVLTSSLSKLPALYLL 121
DB 7 APGLITNFSLATAE-----QCQGE-----TPLENMFLASFYL 39

QY 122 LVFVVGVPANAVTLWMLFFRTRISICT--TVFYTNLAIAADFLFCVTLTPFKIAVHLNGNNW 179
DB 40 LDFILALVGNITLALW-LFIRDHKSQTPANVFLMHLAVADLSCLVLPTRLVYVHFSGNHWP 98
QY 180 FGEVLCRAITTVIFYGNMYCSILLACISINRYLAIVHPPTYRGLPKHTVALVTCGLVWAT 239
DB 99 FGEIACRLTGFLFYLNWYASIIYFLTCISADRFALIVHPVKSUKLRPLVAHLACAFWV 158
QY 240 VFVYMLPFFILKOEYYLVQPDITTDCHDVHNTCESSPPOLY-----YFISLAFEGFLI 292
DB 159 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203
QY 293 PFVLLIYCYAAIIRTLN---LNAVDRWLWYVKASLLILVIFTICFAPSNI---ILIIHA 345
DB 204 PTTTTCVLLIIRSLRQGLRVEKRLKTKAKRMIAIVLAIFLVCFVPYVHNSRVYVHLHYR 263
QY 346 NYYNNYNTDG--LYFYIYIALCLGSLNSCLDPFLYFLMSKTRNHSTAYL 391
DB 264 SHGASCATORILALANRITSCITSLNGALDPIIMYFVFAEKFRHALCNL 311

RESULT 14
US-10-251-385-192

; Sequence 32, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 32
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-32

Query Match 19.0%; Score 410.5; DB 9; Length 339;
Best Local Similarity 30.7%; Pred. No. 6.8e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 62 APNSFEPPFSALEGTGATITVKIKPEESASHLVKNATMGYLTSSLSSTKLIPAIYL 121
DB 7 APPLITNFSLATAE-----QCQGE-----TPLENNLFFASFYL 39
QY 122 LVFVGVGPANAVTLWMLFFTRISICT--TVFYTNLAIADELFCVTLPPFKIAYHLNGNNWV 179
DB 40 LDFILALVGNLTALW-LFIRDHKSCTPANVFLMHLAVADLSVLPVLPTRLVVHFSGNHP 98
QY 180 FGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
DB 99 FGEIACRLTGFLFYLNMYASIVFLTCISADRELAIVHPVKSLLRRPLVAHLACAFWV 158
QY 240 VFYMLPFFILKQEYLVQPDITTCDDVHNTCESSPPFQLY-----YFISLAFGFLI 292
DB 159 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203
QY 293 PFVLIICYAAIIRT-----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
DB 204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263
QY 346 NYNYNTDG--LYFIYIALCLGSLNSCLDPPFLYFLMSKTRNHSTAYL 391
DB 264 SHGASCATORILANRITSCLTSLNGALDPIIMYFFVAEKRHALCNL 311

RESULT 15
US-09-848-889-12
; Sequence 12, Application US/0984889
; Patent No. US20020025555A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Cheng, Muzong
; TITLE OF INVENTION: GPCR DIAGNOSTIC FOR BRAIN CANCER
; FILE REFERENCE: PC-0042 CIP
; CURRENT APPLICATION NUMBER: US/09/848,889
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PERL Program
; SEQ ID NO 12
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020025555A1 g992700
US-09-848-889-12

Query Match 19.0%; Score 410.5; DB 10; Length 339;
Best Local Similarity 30.7%; Pred. No. 6.8e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;
QY 62 APNSFEPPFSALEGTGATITVKIKPEESASHLVKNATMGYLTSSLSSTKLIPAIYL 121
DB 7 APPLITNFSLATAE-----QCQGE-----TPLENNLFFASFYL 39
QY 122 LVFVGVGPANAVTLWMLFFTRISICT--TVFYTNLAIADELFCVTLPPFKIAYHLNGNNWV 179
DB 40 LDFILALVGNLTALW-LFIRDHKSCTPANVFLMHLAVADLSVLPVLPTRLVVHFSGNHP 98
QY 180 FGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
DB 99 FGEIACRLTGFLFYLNMYASIVFLTCISADRELAIVHPVKSLLRRPLVAHLACAFWV 158
QY 240 VFYMLPFFILKQEYLVQPDITTCDDVHNTCESSPPFQLY-----YFISLAFGFLI 292
DB 159 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203
QY 293 PFVLIICYAAIIRT-----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
DB 204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263
QY 346 NYNYNTDG--LYFIYIALCLGSLNSCLDPPFLYFLMSKTRNHSTAYL 391
DB 264 SHGASCATORILANRITSCLTSLNGALDPIIMYFFVAEKRHALCNL 311

RESULT 16
US-09-788-133-2
; Sequence 2, Application US/09788133
; Patent No. US20020052001A1
; GENERAL INFORMATION:
; APPLICANT: GLAXO GROUP LTD
; TITLE OF INVENTION: ASSAY
; FILE REFERENCE: P79011
; CURRENT APPLICATION NUMBER: US/09/788,133
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-788-133-2

Query Match 19.0%; Score 410.5; DB 10; Length 339;
Best Local Similarity 30.7%; Pred. No. 6.8e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 62 APNSFEPPFSALEGTGATITVKIKPEESASHLVKNATMGYLTSSLSSTKLIPAIYL 121
DB 7 APPLITNFSLATAE-----QCQGE-----TPLENNLFFASFYL 39
QY 122 LVFVGVGPANAVTLWMLFFTRISICT--TVFYTNLAIADELFCVTLPPFKIAYHLNGNNWV 179
DB 40 LDFILALVGNLTALW-LFIRDHKSCTPANVFLMHLAVADLSVLPVLPTRLVVHFSGNHP 98
QY 180 FGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYALVTCGLVWAT 239
DB 99 FGEIACRLTGFLFYLNMYASIVFLTCISADRELAIVHPVKSLLRRPLVAHLACAFWV 158
QY 240 VFYMLPFFILKQEYLVQPDITTCDDVHNTCESSPPFQLY-----YFISLAFGFLI 292
DB 159 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 203
QY 293 PFVLIICYAAIIRT-----LNAYDHRWLWYVKASLLILVIFTICFAPSNI---ILIIHHA 345
DB 204 PFITTVTCYLLIIRSLRQGLRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYR 263
QY 346 NYNYNTDG--LYFIYIALCLGSLNSCLDPPFLYFLMSKTRNHSTAYL 391


```
QY 281 YFISLAPFGFLIPFLVLIYYAAI-----IRTL--NAYDHRWLWVYKAS-----LLILVIFTI 331
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 198 --LGACIGVLPPLIILICYSQICCKLFRFAKQNPLETEKSGVKKAKNTIILIVVFL 255
; PRIOR FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
QY 332 CFAPSNILIIH-----HANYNNYNTDGLFYLYL-IALCGLSINSLCDPFLYELMSKTR 384
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
Db 256 CFTPYHVAIIQHMIKKLRFNGLFECQSQRHSFQISLHFTVCLMNFNCMDPFIYFFACK-- 313
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
QY 385 NHSTAYLTXXXNDLREQ 401
; SEQ ID NO 7
Db 314 ----GYKKVYRMKRO 326
; LENGTH: 373
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
US-10-092-135-7

RESULT 20
US-09-828-478-6
; Sequence 6, Application US/09828478
; Patent No. US20020155528A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human CysLT2-Like GPCR
; FILE REFERENCE: 04974.00458
; CURRENT APPLICATION NUMBER: US/09/828,478
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 62 APPNSPEFPFSALEGWGTATITVKIKCPESASHLHVKNATMGYLTSSLSLTKLIPAIYL 121
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 35 APPLCTNFSLTAEE-----QCQGE-----TPLENNMLFASFYL 67
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 122 LVFVGVGPANAVTLMLFFRTRISCT--TVFYTNLAIAIDFLFCVTLPFKIAVHLGNWV 179
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 68 LDFILALVGNLALW-LFIDHKSQTPANVFLMHLAVADLSVILVPLTRLVYHFSGNWHP 126
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 180 FGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHYALVTCGLVWAT 239
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 127 FGEIACRLTGFLFYNMYASIFYLTCISADRFALIVHPVKSLLRPLRYAHLACAFLLVW 186
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 240 VFLYMLPFFLKQEYLVQPDITTCDDVHTNCSSSPFQLY-----YFISLAFTGFLI 292
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 187 VAVAMAPLLVSPQ-----TVQTNHTV--VC-----LQLYREKASHHALVSLA-VAFTF 231
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 293 PFVLIIVCYAAIIRT-----NAYDHRWLWVYKASLLILVIFTICFAPSNI--ILLIHA 345
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 232 PFITTVTCYLLIIRSLRQGRVSEKRLTKAVRMIAIVLALFLVCVFPYHVRNRSVYLHYR 291
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 346 NYYYNNTDGL--LVFYIYIALCGLSINSLCDPFLYELMSKTRNHSTAYL 391
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 292 SHGASCATORILALANRITSCLTSLNGALDPINYYFFVAEKFRALCNL 339
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

RESULT 21
US-10-092-135-7
; Sequence 7, Application US/10092135
; Publication No. US20030054374A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
```

```
; TITLE OF INVENTION: HGPBRMY27
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/273,808
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 60/278,983
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 373
; TYPE: PRT
; ORGANISM: RATTUS NORVEGICUS
US-10-092-135-7

Query Match 19.0%; Score 409.5; DB 9; Length 373;
Best Local Similarity 28.9%; Pred. No. 9.1e-26;
Matches 111; Conservative 64; Mismatches 130; Indels 79; Gaps 13;

QY 69 EFPFSALEG-----WTGATITVKIKCPESASHLHVKNATMGYLTSSLSLTKL 115
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 3 EYPWSAVPNGTDAAFLAGLGLWGNSTI-----ASTAAVSSSFRCL 44
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 116 I-----PAIYLLVFGVGPANAVTLMLFFRTRISCT--TVFYTNLAIAIDFLFCVTL 165
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 45 IKTGQFYLYPAVYILVFIIGLNSVAIWMFVFMKPMWSGISVYMFNLALADFLVLT 104
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 166 PKIAYHLGNWVGEVLCRATTVIFYGNMYCSILLACISINRYLAIVHPFTYRGLPK 225
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 105 PALIFYFNKTDWIFGDVNMCKLQRFHFVNLYGSILFLTCISAHRYSGVYVPLKSLGRKL 164
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 226 HTVALVTCGLVWATVFLYMLPFFLKQEY-----LVQPDITTCDDVHTNCSSSPFQLY 281
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 165 KNAIYVSVLWLVVVAISPIL-----FYSGTGRKNKTVICYD-----STSDEYLSY 214
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 282 FI-----SLAFFGLIPFLVLIYYCYAAIIRTNAVD-----HRWLWYKASLLILVIF 329
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 215 FIVSMCTTVAM--FCIPLVLILGCGYLIVRALYIKDLNPLRRKRSIYLV--IIVLTVF 269
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-828-478-6

Query Match 19.0%; Score 410.5; DB 9; Length 367;
Best Local Similarity 30.7%; Pred. No. 7.4e-26;
Matches 107; Conservative 48; Mismatches 132; Indels 61; Gaps 12;

QY 330 TFCFAPSNIILIH-HANYYYNN-----TDGLFYIYIALCGLSINSLCDPFLYELMSK 382
; FILE REFERENCE: DOI134.NP
; CURRENT APPLICATION NUMBER: US/10/092,135
Db 270 AVSYIPFHYMKTMLNRARLDFTPEMCDNFDRVYATYQVTRGLASLNSCVDPIYFLAGD 329
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,196
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/254,876
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 17
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-937A-17

RESULT 22
US-09-827-937A-17
; Sequence 17, Application US/09827937A
; Patent No. US20020052043A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yi
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human G-Protein Coupled Receptors
; FILE REFERENCE: 1488.1220003
; CURRENT APPLICATION NUMBER: US/09/827,937A
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 08/852,824
; PRIOR FILING DATE: 1997-05-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 17
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-827-937A-17

Query Match 18.8%; Score 405.5; DB 10; Length 348;
Best Local Similarity 30.2%; Pred. No. 1.8e-25;
Matches 99; Conservative 57; Mismatches 115; Indels 57; Gaps 10;
```

QY 106 YLTSLSTKLIPAIYLLVFGVPANAVTLMLFFTRSI-CTTFTYNLAIADEFLFCVT 164
 DB 21 YAHSTARIWPLHYSYLIIGLVGNLLALWIVQNRKKNSTTLSTNLVLSIDILFTTA 80
 QY 165 LPFKIAHLNGNNWVGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTVRLP 224
 DB 81 LPTRIAIYAMGDFWRIGDALCRITALVVFINTYAGVNFMTCLSIDRFAIVHPLRYNKK 140
 QY 225 KHTYALVTCGLVWATVFLYMLPFFI---LKQEYLLVQPDITTCDDVHNTCESS--PFQLY 280
 DB 141 RIEHAKGVCIWFVILVFAOQLPLINPMKQE-----AERITCMEYFNFEETKSLPWIL- 194
 QY 281 YFISLAFPGFLIPFVLIYCAIIL-----RTLNAVDRHRLWLVYK 320
 DB 195 --LGACFVGVPLIIKICYSQICCKLFTAKONPLTEKSGVYKALNTI----- 243
 QY 321 ASLLILVITFCAPSNIILIIH-----HANYYYNNTDGLYFIYL-IALCGLSLSOLD 373
 DB 244 --LLIIVFVLCFTPVHVAIIQHMKKLRFSNFLECSQRHSFOISLHFTVCLMNFNCMD 301
 QY 374 PFLYFLMSKTRNSTAYLTXXNDLREQ 401
 DB 302 PFIYFPACK-----GYKRWMLKQ 323
 RESULT 23
 US-10-092-135-6
 ; Sequence 6, Application US/10092135
 ; Publication No. US20030054374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
 ; FILE REFERENCE: DOI34.NP
 ; CURRENT APPLICATION NUMBER: US/10/092,135
 ; PRIOR FILING DATE: 2002-03-06
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 6
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: HOMO SAPIENS
 ; US-10-092-135-6

Query Match 18.5%; Score 399; DB 9; Length 373;
 Best Local Similarity 29.7%; Pred. No. 6.6e-25;
 Matches 113; Conservative 66; Mismatches 143; Indels 58; Gaps 14;
 QY 52 PTLPIKT---FRGAPPNSFEFFPSALEGTGATITVKIKCPESASHLHVKNATMGYLT 108
 DB 7 PAVPNGTDAFLAGPGSS-----NGNSTV-----ASTAAVSSSFKALT 45
 QY 109 -SSLSTKLIPAIYLLVFGVPANAVTLMLFFTRSI-CTTFTYNLAIADEFLFCVTLP 166
 DB 46 KTGFQFYLLPAVYILVFIIGLVNSVAIMWFVHMKPWSGISVYMFNLALADFLYVLTLP 105
 QY 167 FKAYHLNGNNWVGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTVRLPKH 226
 DB 106 ALIIFYFNKTDWIFGDAMCKLQRFIHNLYGSLIFLTCSIAHRYSGVYVPLKSLGLKK 165
 QY 227 TYALVTCGLVWATVFLYMLPFFILKQBYLVQPDIT--TCHDVHNTCESSPFQLYYFI-- 283
 DB 166 KNAICISVLVWLIIVVAISP--ILFYSGTGVRKNKTITCYDT-----TSDEYLRSYFIYS 218
 QY 284 ---SLAFGGLIPFVLIYCAIIRTLNAYD-----HRWLWYKASLLILVITFCIF 333
 DB 219 MCTTVAM--FCVPLVLILGCGYGLIVRALIYKLDLNSPLRRKSIYLV---IIVLTVFAVS 273

QY 334 APSNIILIIH-HANYYYNN-----TDGLYFIYLIALCGLSLSCLDPLFLYFLMSKTRNH 386
 DB 274 ITPHVMKTMNLRLDFOTPAMCAFNDRVYATQVTRGLASLNSCVDPILYFLAGDTFRR 333
 QY 387 STAYLTXXNDLREQQPSQ 406
 DB 334 RLSRATRKASRRSEANLQSK 353
 RESULT 24
 US-10-092-135-5
 ; Sequence 5, Application US/10092135
 ; Publication No. US20030054374A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bristol-Myers Squibb Company
 ; TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR,
 ; FILE REFERENCE: DOI34.NP
 ; CURRENT APPLICATION NUMBER: US/10/092,135
 ; PRIOR FILING DATE: 2002-03-06
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 75
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 5
 ; LENGTH: 373
 ; TYPE: PRT
 ; ORGANISM: BOS TAURUS
 ; US-10-092-135-5

Query Match 18.5%; Score 398; DB 9; Length 373;
 Best Local Similarity 28.5%; Pred. No. 7.9e-25;
 Matches 111; Conservative 66; Mismatches 136; Indels 76; Gaps 15;
 QY 45 DTNNLAKPTLPIKTRGAPPNSFEFFPSALEGTGATI--TVKIKCPESASHLHVKN 102
 DB 14 DTAFLADP-----GSP-----WGNSTVTSTAASPEKCA----- 43
 QY 103 TMGYLT-SSLSTKLIPAIYLLVFGVPANAVTLMLFFTRSI-CTTFTYNLAIADEFL 160
 DB 44 ---LTKTGFQFYLLPAVYILVFIIGLVNSVAIMWFVHMKPWSGISVYMFNLALADFL 99
 QY 161 FCVTLPPFKIAYHLNGNNWVGEVLCRATTIVFYGNMYCSILLACISINRYLAIVHPFTY 220
 DB 100 YVLTLPALIFYFNKTDWIFGDAMCKLQRFIHNLYGSLIFLTCSIAHRYSGVYVPLKS 159
 QY 221 RGLPKHTYALVTCGLVWATVFLYMLPFFILKQBY---LVQPDITTCDDVHNTCESSSP 276
 DB 160 LGRLKKNAVYISVLVWLIWVVGISPII-----FYSGTGIRKNKTITCYDT-----TSDE 209
 QY 277 FQLYYFI-----SLAFGGLIPFVLIYCAIIRTLNAYD-----HRWLWYKASLL 324
 DB 210 YLRSYFIYSGTVMAM--FCVPLVLILGCGYGLIVRALIYKLDLNSPLRRKSIYLV---II 264
 QY 325 ILVITFCAPSNIILIIH-HANYYYNN-----TDGLYFIYLIALCGLSLSCLDPLFLY 377
 DB 265 VLTVFAVSYPHVMKTMNLRLDFOPTMCAFNDRVYATQVTRGLASLNSCVDPILY 324
 QY 378 FLMSKTRNSTAYLTXXNDLREQQPSQ 406
 DB 325 FLAGDTFRRRLSRATRKASRRSEANLQSK 353

RESULT 25
 US-10-024-494-30
 ; Sequence 30, Application US/10024494
 ; Publication No. US20030044898A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LI, YI
 ; CAO, LIANG
 ; ROSEN, CRAIG A.